

# Formal Effects of Informal Labor Supply Evidence from the Syrian refugees in Turkey

Ahmet Gulek\*

July 15, 2023

## Abstract

I study how firms and natives respond to an informal labor supply shock, and what these results imply about our understanding of the informal economy. The key to my analysis is that the overwhelming majority of refugees did not have work permits and therefore could only work informally. Using an IV-DiD design where I instrument for refugee location choice with population-weighted distance from the border, I show that a 1 percentage point (pp) increase in the refugee/native ratio causes not only a 0.23 pp decrease in informal wage employment, but also a 0.13 pp decrease in low-skill formal wage employment among natives. I also find effects on formal firm entry, where I estimate that refugees reduce the number of newly registered small firms, which is consistent with marginal entrepreneurs choosing to remain informal. Using a modification of the canonical labor demand model, I formalize under which conditions an informal labor supply shock can reduce formal employment. My estimates suggest that informal and formal labor are highly substitutable, with an elasticity of substitution of 9. Lastly, I use the model to estimate a policy-relevant counterfactual in which all refugees are provided with work permits. As a benchmark, I predict that if refugees had the same formality rate as the natives, giving work permits would have created 112,000 formal jobs and increased government tax revenue by \$155 million per year.

*Keywords:* Informality, Immigration, Refugee crises , Work permits

---

\*Gulek: PhD student in Economics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA (e-mail: agulek@mit.edu). I am grateful to Daron Acemoglu, Onur Altindag, Josh Angrist, Abhijit Banerjee, Amy Finkelstein, participants in the MIT Labor Lunch and Development Lunch, and seminar participants at Bilkent University for helpful comments. Arif Akkan, Sarah Gao, Simay Kucukkolbasi, and Lucy McMillan provided excellent research assistance. I also acknowledge support from the George and Obie Shultz Fund at MIT.

# 1 Introduction

The last decade has seen a tripling in the number of refugees in the world, from 10 million in 2012 to 34 million in 2022 (UNHCR, 2021). Two aspects differentiate this recent wave of refugee crises from standard immigration episodes. First, 74% of all refugees are hosted by developing countries with sizeable informal sectors. Second, policymakers in host countries face political pressure to protect the interests of their citizens and often withhold work permits from refugees due to fear of negative labor market implications for their already fragile economies.<sup>1</sup> Consequently, refugees constitute an informal labor supply shock, the consequences of which depend on the dynamics between informal and formal sectors.

In this paper, I study how firms and natives respond to an informal labor supply shock, and what these results imply about our understanding of the informal economy. I first show that in the canonical labor demand framework where a representative firm can use both informal and formal labor in production, an informal labor supply shock necessarily reduces natives' wages and employment in the informal sector. However, more informal employment has two competing effects in the formal sector: it makes formal workers more productive because of Q-complementarity, and it also creates competition against formal employees, especially with diminishing returns in labor. Consequently, refugees' effect on the formal sector remains an empirical question. My framework highlights that if informal and formal labor are largely substitutable in production, then an informal labor supply can incentivize firms to be more informal.

To test the implications of this framework, I study the Syrian refugee crisis in Turkey, one of the largest human displacements in recent history. The Syrian civil war displaced nearly 13 million Syrians, 6.6 million of whom sought refuge in neighboring countries. With 3.6 million registered refugees as of 2020, Turkey hosts the largest number of refugees in the world. Turkey is an ideal setting to study for various reasons. First, it is a developing country where 40% of all employment is informal. Second, the overwhelming majority of Syrian refugees in Turkey lack work permits and must seek informal employment. Third, Turkish labor force surveys allow me to observe wages and employment separately for the formal and informal sectors.<sup>2</sup> Fourth, Syrian refugees in Turkey had freedom of movement, which helped generate a labor supply shock that varies in intensity across the country. Together, these facts lend themselves to a quasi-experimental research design to distinguish

---

<sup>1</sup>This fear is apparent in the following quote from the Minister of Work and Social Security of Turkey “There cannot be a general measure to provide [refugees] with work permits because we already have our workforce . . . we are trying to educate and train our unemployed so they can get jobs in Turkey” (Afanasieva, 2015).

<sup>2</sup>By law, employers in Turkey have to pay for the social security coverage of their employees. Hence, the insurance status of a worker determines her formality type.

the direct impact of informal labor supply shocks to the informal sector from their spillovers to the formal sector.

I first study the refugees' impact on native employment in regular, salaried jobs, which I denote as wage employment.<sup>3</sup> For identification, I rely on a shift-share design, where I use an instrument that exploits the empirical fact that travel distance is inversely related to migrant location. Adjusting for pre-trends, I find that a 1 percentage point (pp) increase in the refugee/native ratio decreases native informal wage employment by 0.23 pp, and formal wage employment by 0.13 pp for the low-skilled. The former is predicted by a downward-sloping labor demand curve in the informal sector, but the latter indicates that informal and formal labor are highly substitutable in production. Mapping my empirical results to model parameters, I estimate the elasticity of substitution between formal and informal labor to be around 9. To the best of my knowledge, this is the first paper that estimates this elasticity using a differences-in-differences design.<sup>4</sup> Put differently, this estimate implies that the common modeling assumption of perfect substitutability between informal and formal workers in the recent structural work on the informal sector (Ulyssea, 2018, 2020) is mostly harmless.

I provide several robustness checks to show that these disemployment effects arise from the (informal) labor supply of refugees and not from other confounders that can result in a decrease in labor demand, such as disruptions in trade due to the Syrian Civil War. First, I document no effect on the formal wage employment rate among high-skilled natives. This is a placebo check as Syrian refugees in Turkey are substantially less educated than the Turkish natives and hence are not close substitutes for this sub-population. Second, I show that the native disemployment comes solely from the industries that received larger labor supply shocks. Syrian refugees in Turkey predominantly do not speak Turkish, which limits the sectors they can work in. Survey evidence shows that refugees work more intensely in textile, construction, and agriculture. Consistent with my hypothesis, natives lose salaried jobs only in these sectors. Third, a back-of-the-envelope calculation using refugees' employment rate suggests that the number of low-skilled workers in the economy increased by 3.9%. Consistent with high-skill/low-skill complementarity, the wages of formal, high-skilled workers increase. Fourth, I show that disruptions in trade with Syria were only temporary and were not large enough to change the total export volume from the border regions. Overall, the accumulated

---

<sup>3</sup>Household Labor Force Surveys in Turkey code employment under four categories: wage earners (defined as regular, salaried work), self-employed, unpaid family workers, and employers. Since I study firms' changing incentives to use the available informal labor supply, I mainly focus on wage employment. I provide more details in the Section 3.

<sup>4</sup>The only other work that I could find that estimates this elasticity is Schramm (2014), who exploits changes in tax regimes in Mexico for identification. Still, her evidence comes mostly from time-series variation.

evidence strongly indicates that the labor supply shock of refugees is the main mechanism behind the adverse employment effects in the informal and formal sectors.

I use the model to estimate the labor market impacts of providing refugees with work permits. Most refugees in the world do not have formal labor market access (Clemens et al., 2018); hence understanding the consequences of work permits is of first-order importance for policy. The model highlights a key trade-off for policymakers: work permits increase native employment in the informal sector and decrease it in the formal sector. Furthermore, the decrease in the total informal employment increases informal wages, which then causes firms to demand more formal workers instead, creating more formal jobs in the economy for natives and refugees alike. The magnitude of these changes depends largely on the ratio of refugees that are endowed with formal labor supply. As a benchmark, I predict that if refugees had the same formality rate as the natives, providing work permits would have created 112,000 formal jobs and increased government tax revenue by \$155 million per year. To benchmark this effect, I compare it to how many formal jobs were created in Turkey during a period of economic growth. From 2004 to 2011, Turkey's GDP per capita increased by 87% from \$6,102 to \$11,420; and the informality rate among low-skill wage jobs decreased by 8 pp from 0.45 to 0.37. If all of this decrease in informality can be attributed to economic growth à la La Porta and Shleifer (2014), then providing work permits to refugees would be equivalent to a 15% growth in GDP per capita for creating formal jobs.

I continue by studying how native workers respond to the refugee shock. I find that immigrants *increase* male natives' non-wage employment, primarily self-employment, and unpaid family work, and do not impact females' non-wage employment. This adjustment is of both economic and empirical importance. Economically, the distinction between wage and non-wage employment is interesting because salaried jobs arise partly from firms' labor demand, whereas self-employment is solely a labor supply decision. This result implies that for low-skilled men, the outside option to salaried employment is non-wage employment instead of unemployment.<sup>5</sup> Empirically, the male escape to self-employment *hides* the disemployment effects of the labor supply shock: as men constitute the majority of overall employment, natives' overall employment rate remains statistically unchanged. Only by dividing employment into wage and non-wage components and by analyzing men and women separately do we see the main economic forces at play. This is why focusing on salaried work is crucial to studying firms' changing incentives to hire formal/informal workers.

Lastly, I investigate whether increased informal employment also impacts firms' decision

---

<sup>5</sup>One potential explanation to why men are so attached to employment, i.e., have low reservation wages, is that in the treated regions in Turkey, men are the primary breadwinner of the household. They may be *expected* to keep having some labor market activity to continue providing for their families.

to register. By utilizing different datasets on formal firm entry in Turkey, I find that whereas the number of incorporated and/or trading firms (high productivity) increases, the number of non-trading firms (mid productivity) remains unchanged, and the number of small sole proprietorships (low productivity) decreases. I further show that in a model where entrepreneurs can locate in informal and formal sectors, this change in the productivity distribution of new formal firms indicates less productive entrepreneurs choosing to remain unregistered. If true, this would be an additional *informalizing* effect of an informal labor supply shock. However, the lack of credible data sources on unregistered firms in Turkey prevents me from testing whether the number of informal firms has increased. Yet, I conclude that the accumulated evidence is highly suggestive.

My quasi-experimental analysis to study the dynamics of informal and formal sectors complements a literature that has studied this topic. Much of this early work was theoretical (Rauch, 1991; Amaral and Quintin, 2006), and most of the recent work has been based on calibrating/estimating structural models (Bosch and Esteban-Pretel, 2012; Meghir et al., 2015; Ulyssea, 2018). A notable exception is Delgado-Prieto (2021), who studies the labor market consequences of the Venezuelan refugee shock in Colombia. He finds negative employment effects in the formal sector but no employment effects in the informal sector, which he rationalizes via a partial equilibrium model inspired by Ulyssea (2018). However, his analysis does not separate salaried and non-salaried employment and does not focus on the role of work permits. As I show in this paper, low-skill men's propensity to transition to non-salaried jobs can hide important economic adjustments in the labor markets. His approach can thus be seen as complementary to the one proposed here in this paper, which focuses on how both firms and natives respond to an informal labor supply shock, and the role of work permits in explaining these effects.

My counterfactual prediction on the formalizing effects of work permits is also related to a literature that studies the impact of different formalization policies in developing countries (Monteiro and Assunção, 2012; De Andrade et al., 2016; Rocha et al., 2018). Most similar to my setting, there are two papers that focus on the value of work permits in refugee crises. On the policy front, Clemens et al. (2018) provide economic arguments as to why providing work permits to refugees can substantially benefit refugees and natives alike. Empirically, Bahar et al. (2021) study the effects of giving Venezuelan refugees work permits. My contribution to this literature is documenting that not providing work permits to refugees acts as an informalizing incentive for firms.

This paper is related to a large literature studying the effects of immigration using refugee shocks. Examples of such episodes include the Mariel Boatlift (Card, 1990), the Algerian war of independence (Hunt, 1992), Jewish emigres to Israel (Friedberg, 2001), the Yugoslav

wars (Angrist and Kugler, 2003), and the Venezuelan refugee crisis Lebow (2022). Despite 30 years of work, whether immigrants cause native disemployment is still debated (Borjas and Monras, 2017; Peri and Yasenov, 2019). Several factors distinguish my setting from the existing literature. First, the treated Turkish regions received substantially more immigrants per native than the aforementioned studies. For example, Mariel Boatlift had increased the labor force of Miami by 8%. In comparison, one city in Turkey (Kilis) observed an increase of 94%. Second, I show the importance of distinguishing between different types of work while studying the impact on natives' labor market outcomes. My findings imply that the labor market effects of immigrants in economies where self-employment is a viable option can be more nuanced than the effect on overall employment and wages.

More recently, several papers investigated the effects of the Syrian refugees on the Turkish labor markets (Del Carpio and Wagner, 2015; Tumen, 2016; Ceritoglu et al., 2017; Akgündüz and Torun, 2020; Erten and Keskin, 2021; Aksu et al., 2022; Cengiz and Tekgüç, 2022; Demirci and Kırdar, 2023) and on firm entry (Altındağ et al., 2020; Akgunduz et al., 2022).<sup>6</sup> Using different identification strategies, this literature mostly found confounding results. I contribute to this literature in several dimensions. First, I show that these opposing findings on native employment result from a combination of (1) not separating employment into components that are governed by different economic forces (mainly wage and non-wage employment) and (2) not adjusting for pre-trends in the shift-share design. Making these economic and econometric adjustments reveals the disemployment effect of natives in both the informal and formal sectors. Second, I provide a tractable model of informal and formal labor demand that both rationalizes these findings and isolates the relevant economic forces in the short run. Future work can build on these economic forces to study the long-run implications of this large labor supply shock. Third, on firm entry, I show that the change in the type of new entrants is *suggestive* of smaller/less productive firms choosing to remain informal. Future research can investigate this effect in more depth by collecting data on informal firms in Turkey.<sup>7</sup>

The rest of this paper is structured as follows. Section 2 provides some background on the Syrian refugee crisis, Section 3 summarizes the data used in the empirical setting, Section 4 explains the identification strategy, Section 5 shows the empirical results on native employment, Section 6 shows the empirical results on formal firm entry, Section 7 develops the model, and Section 8 concludes.

---

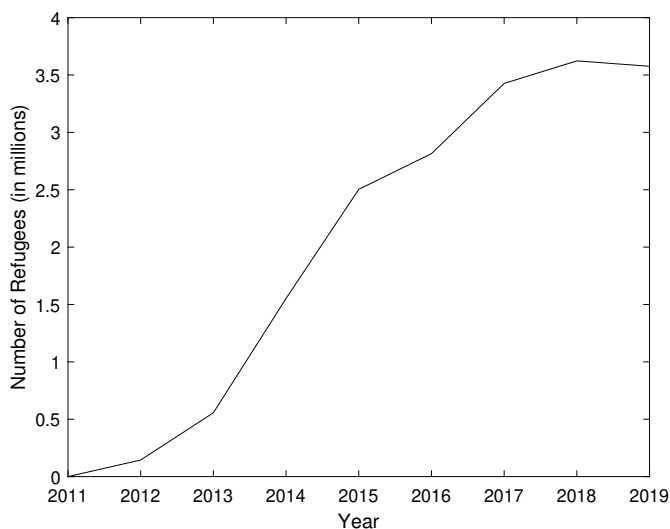
<sup>6</sup>I provide an in-depth discussion of the confounding results and the problems with the identification strategies in the Appendix Section C.

<sup>7</sup>Ozar (2003) is the only rigorous data collection effort on informal firms in Turkey. She finds that around 4% of firms self-declare that they are not registered. Updating this study today, where millions of refugees cannot work formally, could reveal some interesting results. I leave this for future work.

## 2 Background

The Syrian Civil War started in March 2011. By 2017, 6 million Syrians had sought shelter outside of Syria, primarily in the neighboring countries Turkey, Lebanon, Jordan, and Iraq. With 3.47 million registered Syrian refugees, Turkey hosts the highest number of refugees in the world (UNHCR, 2022). The first waves of refugees began arriving in Turkey in late 2011, and the numbers remained small until mid-2012 (İçduygu, 2015). As the violent clashes intensified in the following months, there was a substantial increase in Syrians seeking refuge at the Turkish-Syrian border. Figure 1 shows how the number of Syrian refugees in Turkey has evolved over time. There were around 170 thousand refugees by 2012, 500 thousand by 2013, 1.6 million by 2014, 2.5 million by 2015, and close to 3 million by 2016.

Figure 1: Timeseries of the number of Syrian refugees in Turkey



The Turkish government initially tried to host the Syrians in 25 refugee camps in the southern part of the country across the Turkish-Syrian border. However, these camps quickly exceeded capacity as the number of arriving refugees increased. The refugees thus dispersed across Turkey in heterogeneous quantities.<sup>8</sup> The Figure 2 shows the distribution of the number of Syrian refugees per 100 natives in Turkey at the province level.<sup>9</sup> Refugees are more densely located in regions closer to the border. In fact, distance to the populous governorates in Syria is a strong predictor of refugee settlement. I use this information as part of my identification strategy.

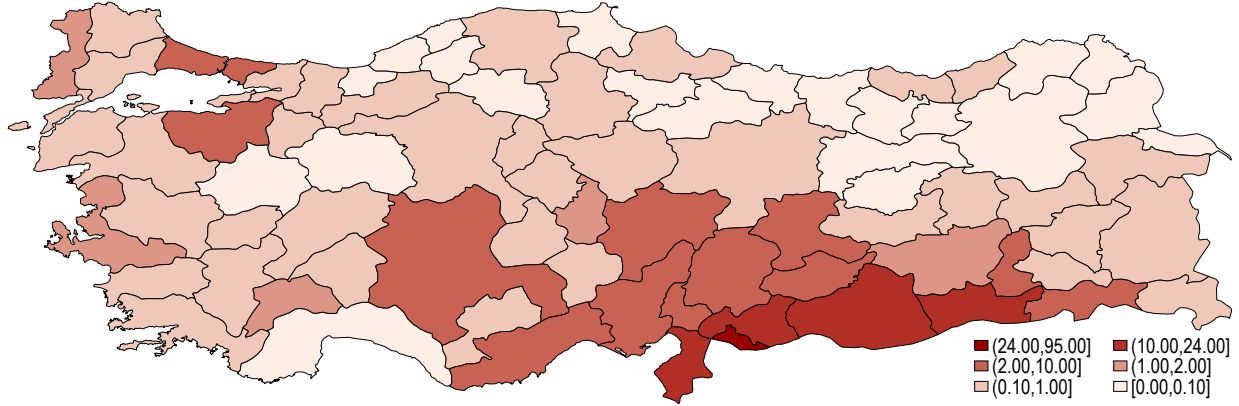
A great majority of Syrians came under the temporary protection category, which permits

---

<sup>8</sup>By 2017, only 8% of the refugees lived inside the camps.

<sup>9</sup>Turkey does not share the education and age break-down of refugees at the province level, which prevents me from exploiting that variation.

Figure 2: Share of Syrian refugees in Turkish population (in%) in 2015



access to health care, education and freedom of movement.<sup>10</sup> Since the temporary protection regime does not offer work authorization, the vast majority of the Syrian labor force works in the informal sector.<sup>11</sup> By the end of 2015, the end of the time period of this study, around 7,300 work permits were issued for Syrian refugees in Turkey.

Syrian refugees are disproportionately less educated compared to Turkish natives.<sup>12</sup> Table 1 compares the education levels of Syrian refugees in Turkey with that of Turkish natives. For instance, 21% of Syrian refugees did not finish primary school compared to 12% of Turkish natives. In addition, 83% of Syrian refugees do not have a high school degree, in contrast to the 61% of Turkish natives. Taking into account the potential educational downgrading and the fact that most Syrian refugees have only basic Turkish language skills, the Syrian refugee shock can be interpreted as a low-skill labor supply shock for the Turkish labor markets.

There is no representative survey on Syrian refugees' employment outcomes. Labor force surveys conducted by the Turkish Statistical Institute do not sample from refugees. The only data we have comes from randomized surveys conducted on ESSN applicants by the Turkish Red Crescent and WFP. This complicates the interpretability of these estimates. Nonetheless, they shed some light on the employment outcomes of natives. Here, I provide a short summary of the findings of these surveys that are relevant for this paper.

According to these surveys, refugees have an astonishing 84% employment rate as opposed

<sup>10</sup>In technical terms, the Syrian population who fled to Turkey are given the temporary protection status, which is different than the full refugee status defined by the Geneva Convention for Refugees. UNHCR uses "refugee-like" term to encapsulate the various forms of protection across countries. I adopt this terminology in line with the literature.

<sup>11</sup>Turkey has passed a law in 2016 to ease the process of acquiring work permits for Syrians. However, the take-up was minimal, potentially due to existing frictions. As of March 2019, only 31,000 Syrian refugees (1.5% of the working-age Syrians) had work permits.

<sup>12</sup>This is due to two reasons. First, Syria was less developed than Turkey, hence it had a lower-educated workforce. Second, highly educated Syrians were more likely to go to Europe.



Table 1: Educational Attainment of Syrian refugees and Natives

Educational Attainment	Syrian migrants (age 18+)	Natives (Age: 18-64)
No degree	0.21	0.12
Primary school	0.42	0.33
Secondary school	0.20	0.16
High school	0.10	0.20
Some college and above	0.08	0.19

Source: Author’s calculation using 2019 Household Labor Force Survey for natives, and Turkish Red Crescent and WFP (2019) for the Syrian refugees.

to 51% for Turkish natives (Turkish Red Crescent and WFP, 2019). The employment rates are high for both men (87%) and women (68%). In contrast, only 68% of native men and 29% of native women are employed. The high employment rates of refugees can be explained by the limited capital they brought to Turkey, and hence lower reservation wages. Refugees have comparative disadvantage in industries that require language skills as only 3% are proficient in Turkish. Perhaps not surprisingly, refugees work mostly in textile (19%), but also in construction (12%) and agriculture (10%). 47% of employed refugees work in regular jobs, defined as a job with a fixed salary and working hours. Textile also has the highest share of refugees in regular positions as 79% of the textile workers have regular positions. The average monthly income of refugees was 1058 TRY in 2019. In contrast, natives in the informal sector made on average 1565 per month in the same year.

### 3 Data

This study combines several administrative, survey, and public datasets. In this section, I introduce the main datasets and explain their properties.

#### Labor Market Outcomes

Information about the labor market outcomes of the Turkish natives comes from the 2004–2016 Turkish Household Labor Force Surveys (HLFS) conducted by the Turkish Statistical Institute (TurkSTAT). They provide detailed information about demographics and labor market outcomes. HLFS is representative at the NUTS-2 level, which consists of 26 regions.<sup>13</sup>

<sup>13</sup>TurkSTAT follows the three levels of NUTS, Nomenclature of Territorial Units for Statistics, defined by the European Union. Under the NUTS definition, Turkey is divided into 11 NUTS-1 regions, 26 NUTS-2 regions, and 81 Nuts-3 regions (province level). I perform all of the analyses in this paper at the 26 NUTS-2 level to keep consistency across different datasets.

The sampling is based on the national address database and does not cover the Syrian refugees that are under temporary protection.

HLFS codes employment under four categories. Between 2004–2016, 60% of employed natives were regular salaried workers (which I denote as wage-earners), 21% were self-employed, 14% were unpaid family workers, and 5% were employers. I combine the latter three under one “non-wage employment” category.<sup>14</sup> This allows me to distinguish between jobs that are partly determined by the labor demand of firms from jobs that depend solely on individual labor supply decision. This distinction is critical in understanding how firms’ react to the informal labor supply shock. For instance, consider a native who loses his formal, salaried job due to being replaced by refugees. This native may keep “working” as an unpaid family worker or trade items at the local markets as a self-employed person. The latter can also be a formal work if the worker pays his social security benefits. Either way, this native would appear as “employed” under the HLFS, even though he was replaced by his employer. To be able to observe this type of transitions, I study wage employment and non-wage employment separately, while focusing on wage employment as the key outcome of interest.<sup>15</sup>

I distinguish between formal and informal employment through a question about the social security coverage. This question is: “Does your job provide any social coverage?” By law, employers in Turkey must provide social insurance coverage for their workers. Consequently, all formal workers are insured and no informal worker can be insured by the employer. Hence, assuming that workers report truthfully in the surveys, I can observe wages and employment in both formal and informal sectors. Although self-reported, insurance status is arguably a good predictor of formality for two reasons. First, there is no incentive for workers to misreport their insurance status. It is not illegal to *work* informally, it is only illegal to *employ* informally. Second, the descriptive statistics on formal and informal employment using insurance status are consistent with the general knowledge on informality. Across regions and industries, the informality rate (defined by the ratio of employment that is informal) decreases by education, is higher in less-developed regions and in industries like Agriculture, Construction, and Textile that are known to rely on informal labor.

I divide the natives into three informality-skill categories: informal employment, low-skill formal employment (those that do not have a high school degree) and high-skill formal employment (those that have a high-school degree). I estimate the impact of Syrian refugees on these 3 groups separately. I do not distinguish between low-skill and high-skill informal natives because there are few high-skill informal natives in Turkey, and they are most likely

---

<sup>14</sup>In general, wage jobs are more desirable than non-wage jobs. Not surprisingly, the probability of a job being a salaried job increases with education, formality, and local GDP.

<sup>15</sup>Furthermore, HLFS collects income information only on wage-earners. Naturally, this also provides an easier comparison between the results on wages and employment as the information comes from the same.

negatively selected. Since overwhelming majority of the Syrian refugees in Turkey did not finish high-school in the Turkish standards, they are predominantly a low-skill informal labor supply. Therefore, they are a closer substitute for low-skill natives.

The wage employment and non-wage employment statistics among different types of natives and industries can be found on Table A.1 in the Appendix. In the aggregate, 49% of prime-age men are employed in salaried jobs, and 25% are employed in non-wage jobs. These numbers fall to 16% and 13%, respectively, for women. Across industries, services are the largest component of wage-employment, and agriculture is the largest component of non-wage employment.

The fact that Turkstat does not sample from refugees makes it difficult to provide the same employment statistics for refugees. However, the propensity of refugees' working in salaried jobs vs non-salaried jobs is important for the interpretation of my results. Turkish Red Crescent and WFP (2019) find that in 2019 refugees had 84% employment rate; and among the employed, 47% were working in regular jobs with fixed salary and working hours. This is more restrictive than the wage employment definition used by the Turkstat, so the wage employment rate of refugees should be even higher.<sup>16</sup> Moreover, the employment rates were likely even higher before the ESSN began due to the income effect of the cash transfer. Taking all of these stats into account, I assume that for every 100 Syrians that arrived to Turkey, 45 were working as wage-employed, which is a conservative estimate.

## Firm entry

To study the extensive margin adjustment of firms, I leverage data on firm formation (i.e., formal firm entry) from three different sources. First, Union of Chambers and Commodity Exchanges of Turkey (TOBB, in Turkish) publishes the number of incorporated firms in Turkey since 2010. This data covers the incorporated new firms (tacir), but does not include sole proprietorships (esnaf). The latter is covered in the Annual Business Registers Framework (Yıllık İş Kayıtları Çerçevesi) of Turkstat, which accounts for the universe of formal firms in Turkey since 2009. The difference between the two types of firms is related to the industry of operation and income. In general, sole proprietorships are smaller in magnitude, and hence more susceptible for informality in theory. Third, I use the data from the Entrepreneur Information System of the Ministry of Industry and Technology (GBS), which also covers the universe of formal firms like Turkstat, but further allows me to separate firms that participate in international trade. On an average year, there are 109 thousand new incorporated firms in Turkey. The average number of new formal firms (including small

---

<sup>16</sup>For example, most work in construction is salaried but irregular.

esnafs) is around 350 thousand in Turkstat and 304 thousand at GBS.<sup>17</sup> Of these firms GBS, 8.7 thousand export and 9.1 thousand import at least once in their lifetime.

Turkish institutions do not collect data on informal firms. Ozar (2003) is the only rigorous data collection effort on informal firms in Turkey. She finds that around 4% of firms self-declare that they are not registered. The actual number is likely even higher because being informal as a firm is illegal. Consequently, informal firms have incentives to either not be interviewed, or even lying conditional on being interviewed. Put differently, 4% firm informality is arguably very little for a country with 40% labor informality. As a comparison, Turkey and Brazil had similar GDP per capita and informal employment rates (40% and 46%, respectively) in 2011. Yet, extensive margin informality of firms with less than 5 employees in Brazil is above 30% (Ulyssea, 2018).<sup>1819</sup>

## Syrian Refugees

The data on Syrian refugees in Turkey comes from a few different sources. In the main text, I rely on the data from the Directorate Generale of Migration Management of Turkey, which provides the provincial distribution of refugees starting from 2012. There is some uncertainty about the exact distribution of refugees across Turkey, especially between 2011 and 2012. However, since most of this continuous treatment occurred after 2013 (for which we have more reliable data), the estimates remain robust to using different data sources for 2011 and 2012.

## Additional Data sources

For robustness checks, I rely on a number of additional data sources. I gathered the foreign trade statistics at the province-country level from the Foreign Trade Statistics Micro Data Set of Turkstat. The trade data allows me to control for trade shocks from the Syrian War in some of the robustness checks in the Appendix. Lastly, I also use the provincial electricity consumption data from Turkstat to proxy for firm activity.

---

<sup>17</sup>Turktstat and GBS data do not exactly match, which is due to the different administrative sources they draw the data from. However, my qualitative results remain robust to using either data source.

<sup>18</sup>It is also worth mentioning that 4% of firms being informal is an equilibrium outcome. If new informal firms have higher exit probabilities than new formal firms, than the ratio of informal firms among new firms would be higher. For exmaple, Ulyssea (2018) estimates that the informal exit probability is 3 times of that of formal firms in Brazil. If this ratio is similar in Turkey, this would imply that at least 12% of new firms in Turkey in a given year remain informal.

<sup>19</sup>It is worth pointing out that forming a business in Turkey is easier than in Brazil. According to World Bank's "Doing Business" report in 2011, Turkey ranked 65/183 in "ease of doing business", whereas Brazil ranked 127 (World Bank, 2010). Since the difficulty in doing business is associated with firm informality, we would expect to have higher informality rates in Brazil than in Turkey.

## 4 Identification

My identification strategy exploits the differential intensity of Syrian refugees across space and time. The unit of analysis is region-year. I define the treatment  $R_{p,t}$  as the ratio of the number of refugees to natives in region  $p$  and year  $t$ . The outcome variables are statistics of the local labor market conditions, such as native employment rates. If the local labor market conditions are a determinant of refugee settlement, then a simple difference in differences strategy would give biased estimates.

To circumvent this bias, I exploit the fact that travel distance is a strong predictor of migrant settlement in forced migration episodes (Angrist and Kugler, 2003; Del Carpio and Wagner, 2015). In my setting, Syrian refugees have a strong tendency to settle closer to the border. Moreover, given the same distance to the nearest bordercrossing, there are more Syrian refugees in Turkish provinces that are closer to the more populous Syrian governorates. I use a weighted-distance instrument  $Z_p$  that exploits these two facts to predict refugee settlement:

$$Z_p = \sum_{s=1}^{13} \lambda_s \frac{1}{d_{p,s}} \quad (1)$$

where  $d_{p,s}$  is the travel distance (in km) between Turkish region  $p$  and Syrian governorate  $s$ , and  $\lambda_s$  is the weight given to Syrian governorate  $s$ .<sup>20</sup> Different weights of  $\lambda$  have been used in the literature. In practice, weights matter little. In the main text, I rely on the weights suggested by Aksu et al. (2022).

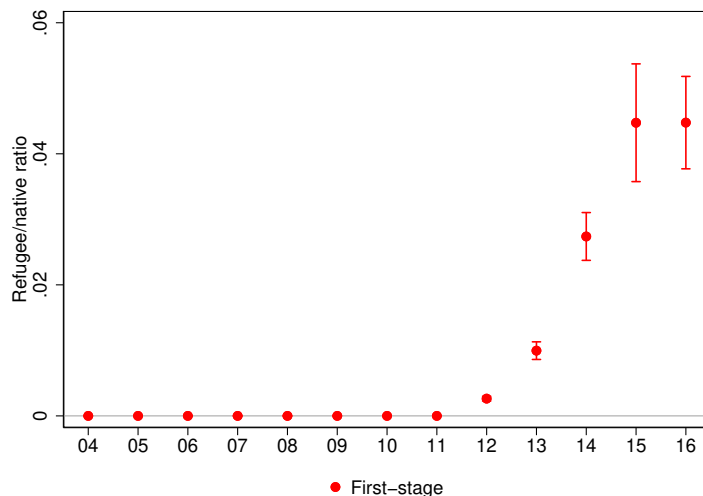
$$\lambda_s = \underbrace{\frac{\frac{1}{d_{s,T}}}{\frac{1}{d_{s,T}} + \frac{1}{d_{s,L}} + \frac{1}{d_{s,J}} + \frac{1}{d_{s,I}}}}_{\substack{\text{Relative distance} \\ \text{to Turkey}}} \times \underbrace{\pi_s}_{\substack{\text{Pop.} \\ \text{share}}} \quad (2)$$

where  $d_{s,c}$   $c \in \{T, L, J, I\}$  is the travel distance between Syrian region  $s$  to Turkey, Lebanon, Jordan, Iraq respectively; and  $\pi_s$  is the population share in 2011, which I calculate using the 2011 census undertaken by the Central Bureau of Statistics of Syria. The intuition behind this set of weights is that all else equal, more refugees leave Syria from the populous regions, and refugees are more likely to come to Turkey if they were closer to Turkey than the other neighboring countries. Put differently, the instrument *predicts* the governorate-origins of the Syrian refugees in Turkey.<sup>21</sup>

<sup>20</sup>I used the city centers in each region to calculate the travel distance. The data is available upon request.

<sup>21</sup>I show how these predictions compare to actual numbers in the Appendix Table A.2.

Figure 3: Event Study of the First-stage



Notes: The regression equation is:  $R_{p,t} = \sum_{j \neq 2009} \theta_j(\text{year}_j \times \tilde{Z}_p) + f_p + f_t + \eta_{p,t}$ , where  $\tilde{Z}_p$  is the standardized version of the instrument  $Z_p$  to have economically meaningful coefficients,  $f_p$  and  $f_t$  are region and year fixed effects. Standard errors are clustered at the nuts2 region level. The 95% confidence interval is shown.

The DiD design and the shift-share nature of the instrument allows me to estimate both the first-stage and the reduced-form as an event-study. To show the treatment intensity predicted by the instrument, I estimate the following event study design.

$$R_{p,t} = \sum_{j \neq 2009} \theta_j(\text{year}_j \times \tilde{Z}_p) + f_p + f_t + \eta_{p,t} \quad (3)$$

where  $\tilde{Z}_p$  is the standardized version of the instrument  $Z_p$  to have economically meaningful coefficients,  $f_p$  and  $f_t$  are region and year fixed effects. I cluster the standard errors at the nuts2-region level. Recall that there are no refugees in Turkey before 2011, so  $\theta_j = 0$  if  $j < 2012$ . I plot the estimates of  $\theta_j$  in Figure 3. The instrument strongly predicts refugee settlement in all post-years. The joint F-stat for the instrument in years 2012–2016 is 238.

This figure also reveals how the treatment intensity as predicted by the instrument increases overtime. The treatment intensity is low in 2012 as there are fewer refugees. It slightly increases from 2012 to 2013, and increases substantially in 2014 and 2015. This time-series variation is important for identification because given any nonzero effect of refugees on the outcome of interest, we would expect the estimated treatment effect to increase overtime.

The standard identifying assumption of an IV-DiD design with a continuous instrument is that the instrument, distance to border, is orthogonal to local trends. However, this does

not hold for many of the outcomes in my setting. Between 2004–2010 (before the refugee shock begins), regions close to the border observed higher growth in employment rate, wages, and firm entry, leading to a positive trend that is correlated with the instrument.<sup>22</sup>

To make progress, I exploit the empirical fact that pre-trends are approximately linear for most of the outcomes I study throughout the paper. I first estimate region-specific linear trends using data between 2004–2010, and then detrend the outcome by extending the trend post 2010. To be more precise, I define the detrended outcome as:  $\tilde{y}_{p,t} = y_{p,t} - (t - 2004) * \hat{y}_{tr,p}$ , where  $\hat{y}_{tr,p}$  is the estimated linear trend at region  $p$ . I then use the detrended outcome inside the event-study design:

$$\begin{aligned}\tilde{y}_{p,t} &= \sum_{j \neq 2009} \beta_j (\text{year}_j \times \tilde{Z}_p) + \delta_p + \delta_t + \epsilon_{p,t} \\ R_{p,t} &= \sum_{j \neq 2009} \theta_j (\text{year}_j \times \tilde{Z}_p) + f_p + f_t + \eta_{p,t}\end{aligned}\tag{4}$$

where  $\beta$ 's show the standard event-study estimates for the detrended outcome  $\tilde{y}$ , and  $\theta$ 's show the treatment intensity predicted by the instrument. To have economically meaningful reduced-form estimates, I standardize the instrument  $Z_p$  to have mean zero and standard deviation of one. Hence, the coefficients can be interpreted as the effect of having one standard deviation higher iv-weight. The identifying assumption is that distance to border is orthogonal to *deviations* from the region-specific linear trends.<sup>23</sup>

Notice that in the event-study design I estimate the first-stage estimates alongside the reduced-form. This is not standard practice. However, given that the treatment is continuous instead of binary, the evolution of the first-stage is informative. For example, since the refugee treatment intensifies post 2014, we would expect the reduced-form effects to also intensify post 2014. Showing both the first-stage and the reduced-form together allows the researcher to visually check the relationship between the two regression estimates.

After showing the event-study estimates, I also estimate the following IV-design using 2SLS to get economically meaningful estimates:

$$\begin{aligned}\tilde{y}_{p,t} &= \beta R_{p,t} + \delta_p + \delta_t + \epsilon_{p,t} \\ R_{p,t} &= \theta Z_p T_t + f_p + f_t + \eta_{p,t}\end{aligned}\tag{5}$$

---

<sup>22</sup>These pre-trends can be seen in the event-study figures in the Appendix Section B.

<sup>23</sup>Controlling for linear trends via this 2-step procedure has two appealing properties. It estimates the pre-trend in the pre-period, and does not introduce the trend in the first-stage regression. In the Appendix Section C, I show how alternative methodologies that are often used in the literature, such as controlling for trends inside the regression or adding aggregate region-year fixed effects fail to properly account for the pre-trends.

where  $T_t = \mathbb{1}\{year = t\}$  is a year indicator,  $\delta_p, \delta_t, f_p, f_t$  are province and year fixed effects,  $R_{p,t}$  is the refugee/native ratio, and  $\tilde{y}_{p,t}$  is the detrended outcome. This design takes into account both the cross-sectional and the time-series variation of the treatment into account to estimate  $\beta$ . In estimation, I always control for region and time fixed effects, and cluster the standard errors at the region level.

There are a few threats of identification that are worth discussing. Notice that the distance instrument basically compares the regions close to the border with those further away. This comparison may not identify the causal effect of refugees for three main reasons. First, this empirical strategy assumes that the Syrian war’s impact on the Turkish local labor markets, if any, should be orthogonal to distance to border. This can fail if Syria was a major trade partner of regions at the border, and if the war had significantly disrupted the trade flows. Empirically, Syria was not a major trade partner of any region in Turkey. Moreover, even though trade initially fell in 2011 and 2012 at the beginning of the war, it more than recovered at the border regions after 2013. Hence, there was not a significant trade shock that could impact the local labor markets. I provide more details on the evolution of trade flows across regions in the Appendix Figure E.15.

Second, even if refugees impact the regions they go to, given enough time markets could reequilibrate across space through movement of capital and people. This would violate the SUTVA, and the IV-DiD methodology would underestimate the effect of refugees. To prevent this bias I study the short-run. Treatment intensity was economically meaningful only after 2013, and I stop the analysis by 2016.<sup>24</sup> Within this time period, I document only minor changes in the movement of people across space. As I show in Figure D.12, regions that are close to the border faced slightly more out-migration and less in-migration. However, these effects are very low in magnitude. Hence, during the time of study, there is no evidence for an economically meaningful change in the movement of people across regions that can bias my results. Therefore, I don’t think violation of SUTVA is a first-order concern.

Lastly, my solution to pre-trends is fitting a linear line in the pre-period and extending it to the post. For identification, deviations from the pre-trend should be orthogonal to the instrument. This assumption is likely to fail in the long run. Any “catching-up” effect between the less developed south-east regions and the rest of Turkey has to slow down eventually. Restricting the analysis to the short run also helps limit this bias.

---

<sup>24</sup>Other reasons why I stop the analysis by 2016 include a minimum wage increase in 2016, and the beginning of the Emergency Social Safety Net (ESSN) program in which refugees were given relatively large cash transfers. Both of these confounders could complicate the interpretability of the estimated effects post 2016.



## 5 Results

### 5.1 Employment

In this section I show refugees' effect on natives' employment outcomes. As I explained in Section 3, I start by isolating the effect on wage-employment as firms' changing incentives to hire labor is captured by regular, salaried workers. I study three types of wage-employment: informal wage-employment, formal low-skill wage-employment and formal high-skill wage-employment. As refugees are predominantly low-skill for the Turkish labor force, we would expect the employment effects to be concentrated within the low-skill natives. I also study the impact on men and women separately.

#### Event-study estimates

I begin by estimating the event-study design shown in equation 4 using detrended data.<sup>25</sup> I plot the estimates in Figure 4. In Panel A, I plot the estimates for informal wage-employment. Across both men and women, controlling for a linear trend eliminates the pre-trend completely. Not only the estimates between 2004-2011 are insignificant, magnitude-wise they are also close to zero. Moreover, the instrument predicts strong disemployment effects. A one standard deviation increase in the instrument is associated with a 1.7 pp decrease in informal wage-employment for men, and a 0.5 pp decrease for women by 2016. The effect for women is smaller in magnitude and statistically insignificant. Also notice that the decrease in employment follows the treatment intensity showed by the first-stage. Meaning, as more refugees arrive, a stronger disemployment effect we observe, which strongly increases the credibility of these results.

In Panel B, I plot the estimates for formal wage-employment rates among low-skill natives. Similar to the informal employment, the instrument is orthogonal to deviations from the linear trend in the pre-period: between 2004–2011, the instrument does not predict a meaningful change in employment. In the post period, even though refugees cannot work formally, a one standard deviation increase in the instrument is associated with a 0.7 pp decrease in formal employment for low-skill men, and a similar 0.65 pp decrease for women by 2016. Similarly to informal employment, the reduced form effects become apparent starting from 2014 as the treatment intensity increases. In Panel C, I plot the estimates for formal wage-employment rates among high-skill natives. The estimates in the post period are all insignificant and most are close to zero in magnitude, both for women and men.<sup>26</sup>

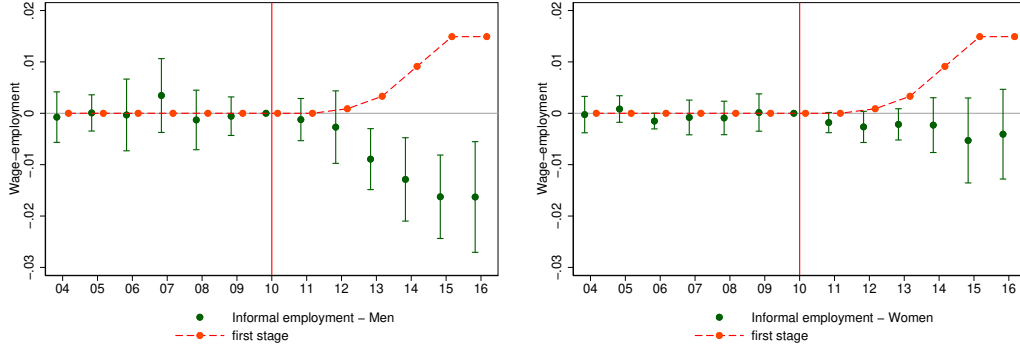
---

<sup>25</sup>The event study figures using raw data can be found in Figure B.1.

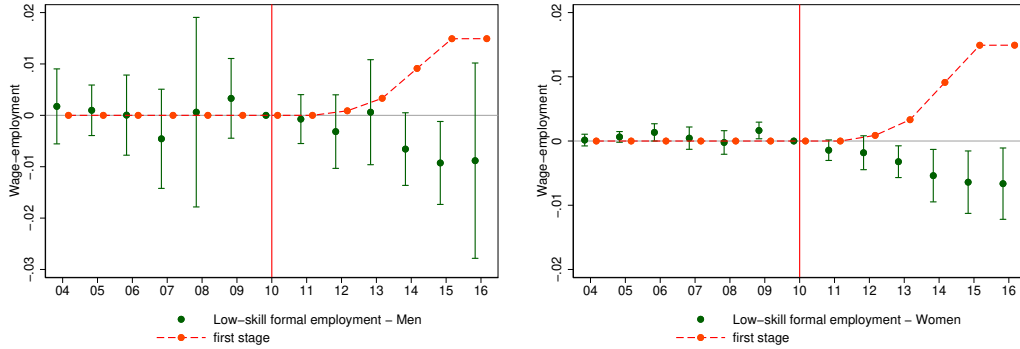
<sup>26</sup>It should be noted that the instrument predicts significant effects in the pre-period for women. This is because the data on high-skill women is more noisy, and the raw data does not show a linear pattern in

Figure 4: Event-study figures with detrended data

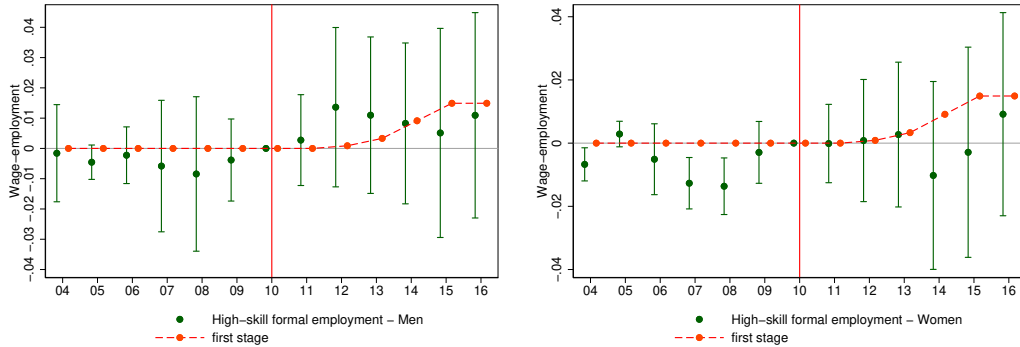
(a) Informal wage-employment



(b) Formal low-skill wage-employment



(c) Formal high-skill wage-employment



Notes: The reduced-form estimates come from the event-study design shown in equation 4. The outcome variable is the informal wage-employment rate in Panel A, formal wage-employment rate for the low-skilled in Panel B, and formal wage-employment rate for the high-skilled in Panel C. The left column shows the estimates for men, and the right column for women. Standard errors are clustered at the region level. The 95% confidence interval is plotted. The red dashed line shows the event-study point estimates for the first-stage, which captures how the treatment intensity predicted by the instrument varies over time.

Taking stock, the event-study figures show that the instrument does not predict economically meaningful effects in the pre-period, which is reassuring for the identification strategy. The instrument is associated with significant decreases in both informal and low-skill formal wage employment rates. The latter is despite refugees' inability to find formal work. These effects become stronger after 2014 as the treatment intensity increases. Lastly, the instrument is not associated with a significant change in high-skill formal employment.<sup>27</sup>

## 2SLS estimates

Event-study figures are suggestive, but they are mostly imprecise because they do not exploit the timeseries variation in treatment intensity. To get economically meaningful and statistically more precise estimates, I estimate equation 5 using 2SLS while clustering standard errors at the region level. I plot the estimates for informal and low-skill formal wage-employment rates for the pooled sample of men and women in the first row of Figure 5. A 1 pp increase in the refugee/native ratio decreases informal wage-employment rate of natives by -0.23 pp, and formal wage-employment rate of low-skill natives by -0.13 pp. Both effects are statistically significant at 0.001 significance level. I also detect null effects on the formal wage-employment rate for high-skill natives, but the imprecision of these estimates render the figure less readable. I show these estimates in Figure E.13 in the Appendix.

To put these numbers into perspective, it is fair to assume that most refugees need to work somehow to survive. Most brought little to no capital, and in the time-period studied there was not a widespread cash or in-kind transfer program targeted for refugees. Even after ESSN started at the end of 2016, employment rates among refugees were high. Among ESSN applicants, a random survey conducted by the Turkish Red Crescent and World Food Program in 2019 found an astonishing 84% employment rate. Among the employed, 47% were working in regular jobs with fixed salary and working hours. This is more restrictive than the wage-employment definition used by the TurkSTAT, so the wage-employment levels of refugees should be even higher. Moreover, due to income effects the employment rates were likely higher before the unconditional cash transfer began. So, I assume that for every 100 Syrians that arrived to Turkey, 45 were working as wage-employed. Consider the following thought experiment. Let region A have 1000 natives in period 1, all low-skill for simplicity. On average, 23.3% of low-skill natives are wage-employed, hence 233 wage-employed natives.

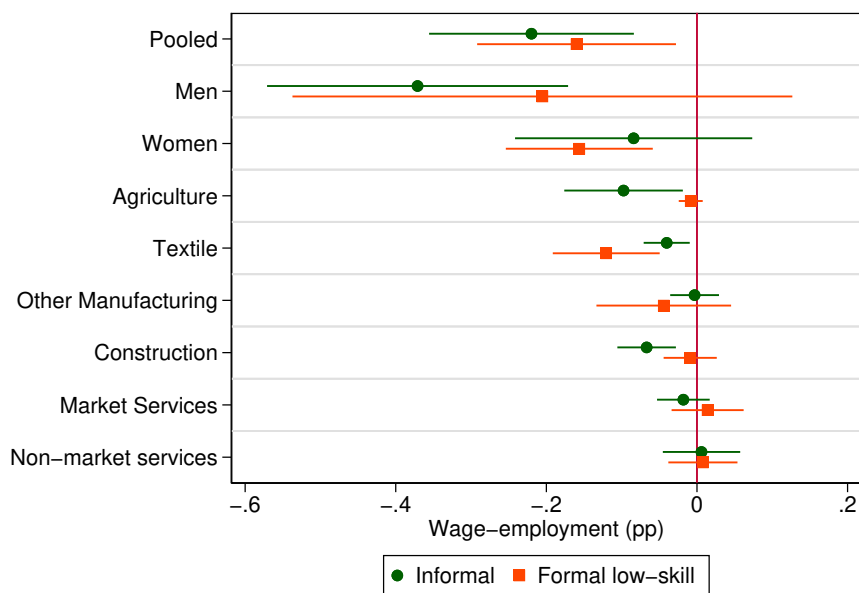
---

the pre-period, which can be seen on Figure B.1. However, both the raw and the detrended data show null effects in 2SLS, so whether a linear trend is a good assumption for this outcome is not a major concern.

<sup>27</sup>It is worth emphasizing that detrending the data is not necessary to find the decline in wage-employment for informal men and formal low-skill women. This is why the earlier literature studying the employment effects of natives were able to detect them by relying on the baseline IV-DiD strategy without adjusting for trends. However, the remaining pieces of the puzzle remained allusive until pre-trends were adjusted for, which is one of my empirical contributions.

In period 2, this region receives 100 refugees, a 10 pp increase in refugee/native ratio. My estimates suggest that this shock leads to 36 natives losing informal and formal wage-jobs combined. In other words, 45 refugees replaced 36 natives in regular, salaried jobs. The total low-skill employment increased by  $9/233 = 3.9\%$ .

Figure 5: Effect of Refugees on native wage-employment rates



Notes: The 2SLS estimates come from estimating equation 5 using the detrended versions of informal wage-employment rate or the formal wage-employment rate for the low-skilled. The first row shows the estimates on the pooled data. The second and third rows condition on men and women separately. Rows 4–9 show the heterogeneity across industries. The industry specifications follow the ISIC standards. Standard errors are clustered at the region level. The 95% confidence intervals are plotted.

In the second and third rows of Figure 5, I plot the estimated effects of refugees on native men and women separately. Among men, a 1 pp increase in refugee/native ratio decreases informal wage-employment by 0.39 pp, and formal low-skill wage-employment by 0.15 pp. The latter effect is imprecisely estimated with a p-value of 0.13. Among women, informal wage-employment falls by 0.09 pp and formal low-skill wage-employment also falls by 0.15 pp (p-value=0.002). One reason why men are facing a larger disemployment in levels (especially for informal work) is that their informal wage-employment rate is higher: 0.11% of all working-age men in Turkey have informal wage-jobs, compared to 3.6% for women. Consequently, the effects are similar in percentages. Perhaps the more interesting finding is that the effects on formal employment for low-skill men and women are similar in levels. This is despite low-skill men being 7 times more likely to be employed formally in a wage-job. Some of this discrepancy is likely due to men and women working in different occupations,

and within occupations, performing different tasks.

Taking stock, these estimates suggest that the informal refugee shock has caused native disemployment in both the informal and formal sectors. Also informative is that only low-skill natives lose jobs in the formal sector, and high-skill natives do not. The fact that low-skill formal natives lose employment and high-skill natives don't is informative about the mechanisms by which informal refugees impact the formal sector. The demand side effects, either via refugees' demanding goods and services, or via product market competition between informal and formal firms, would impact both low-skill and high-skill natives. In contrast, Syrian refugees' labor supply can only substitute low-skill tasks in Turkey. The evidence suggests that formal firms can substitute between formal and informal workers among the low-skilled. Before exploring the implications of these findings further, I investigate the robustness of these estimates.

## 5.2 Supporting Evidence

### Heterogeneity across Industries

Syrian refugees disproportionately work at particular industries due to comparative advantage. Most are not proficient in Turkish, hence they are less likely to perform tasks that require communication in written or spoken language. Consequently, they work predominantly in jobs that require manual work: textile (19%), construction (12%), and agriculture (8%) (Turkish Red Crescent and WFP, 2019). If the disemployment effects of natives are due to the labor supply of Syrian refugees, then we would expect to see higher disemployment effects on the more intensely treated industries.

To test this hypothesis, I separate the native employment into 6 categories: agriculture, textile, other manufacturing, construction, market services, and non-market services following ISIC definitions. Then, I estimate refugees' effect separately on each industry, and plot these effects on rows 4–9 on Figure 5.<sup>28</sup> As expected, the disemployment effects in the informal and formal sectors come solely from these high-intensely treated industries. Informal disemployment occurs mostly in the Agriculture, Construction, and Textile industries. A 1 pp increase in refugee/native ratio decreases informal wage-employment by 1 pp in Agriculture, 0.08 pp in Construction, and 0.04 pp in Textile. In contrast, disemployment in the formal sector (for the low-skilled) occurs mostly in the textile, where a 1 pp increase in refugee/native ratio decreases wage-employment of the low-skilled natives by 10 pp. There

---

<sup>28</sup>The event-study estimates for these outcomes can be found in Figures B.2 and B.3 in the Appendix Section B. Adjusting for a linear trend eliminates the pre-trend for all of the outcomes for which I find an effect.

is also a negative estimate on other manufacturing industries (0.03 pp), but the effect is not statistically significant. Lastly, there is no change in native wage-employment in services.

The finding that refugees do not impact native employment rates in the less intensely treated industries is also interesting because these industries are still *treated*. Absent equilibrium effects, we would still expect to see some native disemployment. There are two additional forces that are working in the opposite direction that can explain this null result. First, refugees demand goods and services, and hence may push the labor demand curve outwards. If the labor supply effect is small enough, it may be completely offset by this demand shock. Second, in an equilibrium model with multiple sectors and flexible native labor supply across sectors, a refugee labor supply shock to one sector would cause the natives to focus their search for jobs in the other sectors, increasing the equilibrium employment rate in these industries. I suspect both effects to play some part in explaining this null effect, but due to data limitations I leave this for future research.

## Wages

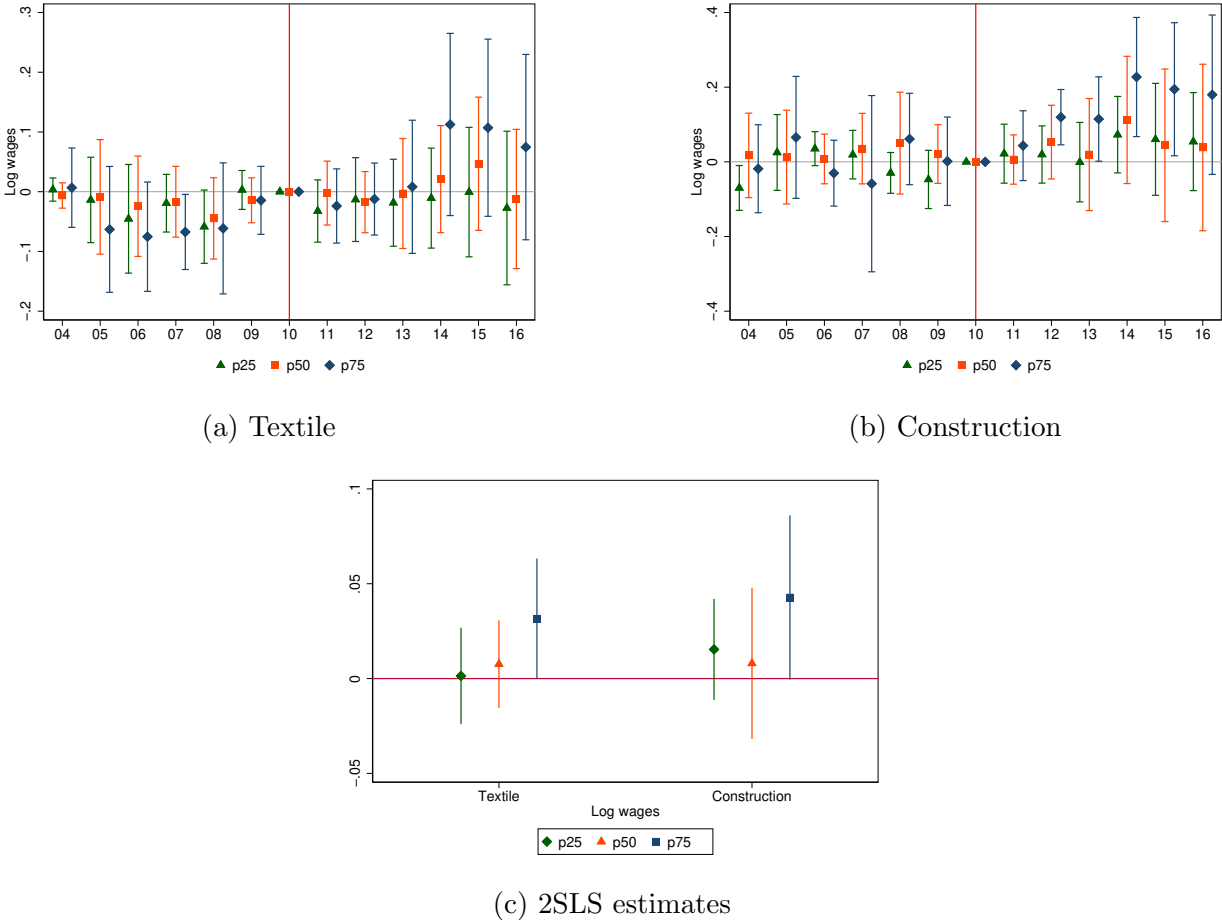
My point estimates, combined with survey evidence on the wage-employment rate of Syrian refugees, imply that the total amount of low-skilled employment in the economy increased. Under the canonical labor demand model, this increase in low-skill labor should increase the productivity and therefore the wages of high-skill labor in the economy. To test this intuition, I estimate refugees' effect on high-skilled natives wages in the formal sector. In particular, I estimate versions of equations 4 and 5, where the outcome variables are the 25th, 50th, and 75th percentiles of wages. I plot the results on Figure 6. For ease of presentation, I show the results only for construction and textile, the two industries that were most heavily impacted by the refugee shock. I omit agriculture because there are practically no formal, high-skilled workers in that industry.

I show the event-study estimates for textile on figure 6a and for construction on figure 6b. In both industries, there are no effects on the 25th and 50th percentiles of the wage distribution. However, there is an increase in the 75th percentile of wages starting from 2014, as the refugee inflow becomes economically meaningful. To test for statistical significance and obtain economically meaningful estimates, I show the 2SLS estimates on figure 6c. I find that a 1 pp increase in refugee/native ratio increases the 75th percentile of the high-skill wage distribution by 3% (p-value:0.51) in textile, and by 4% (p-value: 0.53) in construction.

The fact that high-skill wages in these highly treated industries increase serves multiple purposes. First, assuming low-skill and high-skill complementarity, it provides an additional evidence that the total low-skill employment in the economy increased due to the refugee shock. This was previously predicted by my estimates on native disemployment combined

with survey evidence on refugees' employment rates. Second, it helps eliminate one of the major identification concerns in my design. Regions close to the border could have received negative demand shocks from the Syrian Civil War; e.g., through trade or security channels. Industry specific negative demand shocks would be consistent with my findings on native disemployment. However, negative demand shocks would have decreases wages, not increase them. In short, disemployment of low-skill natives in these industries, together with increases in high-skill wages, can only be explained by the labor supply channel.

Figure 6: Effect of refugees on high-skill formal wages



Notes: The reduced-form estimates come from the event-study design in equation 4, and the 2SLS estimates come from the IV design in equation 5. The outcome variable is the pth percentile of log wages in the textile industry in Panel A, and the pth percentile of log wages in the Construction industry in Panel B. px refers to the xth percentile of the wage distribution. The outcome variables are detrended by fitting a linear trend in the pre-period. Standard errors are clustered at nuts2 region level. The 95% confidence intervals are plotted.

## Non-wage employment

Throughout the empirical investigation thus far, I focused on wage employment instead of overall employment. In this subsection, I show the importance of separating wage employment from non-wage employment while studying the labor demand. This will also shed more light to why my empirical results on native employment are different than the rest of the literature studying the labor market consequences of the Syrian refugees in Turkey.

As I explained in Section 3, there is an economically meaningful distinction between wage-employment and nonwage-employment in Turkey, which can be generalized to similar developing countries. Regular, salaried jobs are jobs in which the employment status of the worker depends on the decision of an employer. If an employer finds cheaper labor that can perform the same tasks, the worker would lose her job. However, anyone who is doing some amount of market activity can correctly declare themselves to be self-employed. For example, when refugees displace natives in the regular jobs in textile, the displaced natives who have strong labor force attachment may still remain *employed* by doing any market activity on their own. In the extreme, the net effect on employment may be zero, even though natives have lost their jobs.

In fact, researchers studying refugees' effect on native employment would not find most of the disemployment effects in the economy, even if they adjusted for pre-trends correctly. To show this, I focus on the the formal employment outcomes of low-skill natives. In particular, I estimate refugees' effect on natives' total employment rate, wage employment rate, and non-wage employment rate separately. I follow the structure in Section ?? and show the heterogeneity in these effects across sex and industry.

I plot the estimates on Figure 7.<sup>29</sup> Looking at the first row, we see that refugees had no effect on natives' total employment. However, this null effect masks a substantial heterogeneity across the employment types. As I showed earlier, natives' formal wage employment decreases considerably. Interestingly, this decline in wage employment is offset by an increase in non-wage employment. This implies that natives who lose their salaried jobs transition into non-salaried market activities.

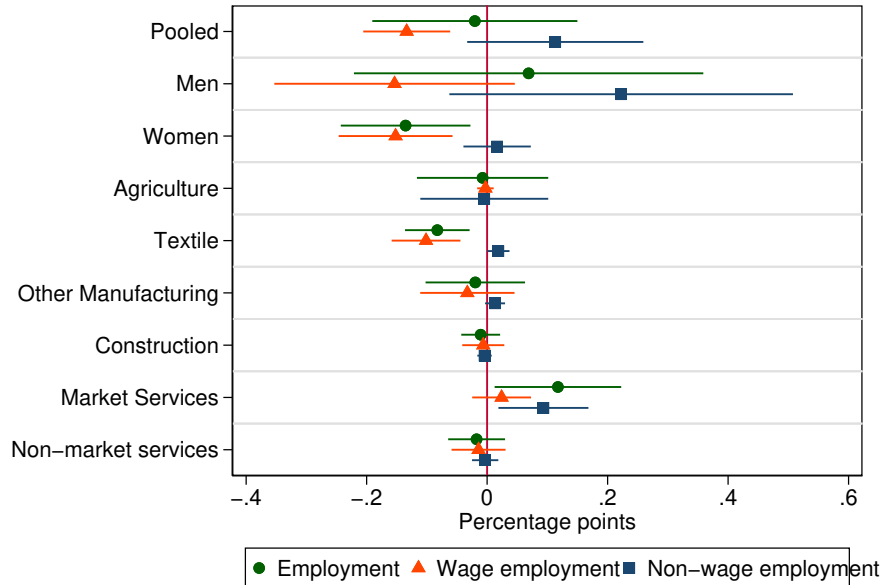
In the second and third rows, I show the heterogeneity across sexes. This exercise reveals that whereas both men and women incur similar decreases in wage employment, it is only the men who are transitioning into non-salaried jobs. My best explanation for this heterogeneity is the men's role as the primary bread-winner of the household in Turkey. Upon losing their primary regular jobs, men end up having to do some type of market activity as their labor

---

<sup>29</sup>The event-study estimates for these outcomes can be found in Figure B.4 in the Appendix Section B. Adjusting for a linear trend eliminates the pre-trend for all of the outcomes for which I find a significant effect.



Figure 7: Refugees' effect on wage and non-wage employment rates



Notes: The 2SLS estimates come from the IV desing in equation 5. The outcome variable is either the informal wage-employment rate or the formal wage-employment rate for the low-skilled. The outcome variables are detrended by fitting a linear trend in the pre-period. The first row shows the estimates on the pooled data. The second and third rows condition on men and women separately. Rows 4–9 show the estimates separately by industry categories. Industries are defined using two-digit NACE codes following ISIC Rev 4 definitions. Standard errors are clustered at nuts2 region level. The 95% confidence intervals are plotted.

force attachment is very high. In contrast, women do not have this cultural preference, and hence do not make this transition upon losing their salaried jobs.

This explanation is also supported by the heterogenous effects across the industries. As I showed earlier regarding wage employment, the disemployment in the formal sector concentrates on the textile sector. In contrast, the increase in non-wage employment comes mostly from the market services. This is very intuitive. Consider the type of self-employed jobs that are easily accessible. It is arguably much harder for a laid-off textile worker to open a textile shop, then to buy and sell goods in the market. Taking stock, this figure shows that both men and women lose their salaried jobs (mostly in textile), but men transition into non-salaried jobs in market services as they *have to* remain attached to the labor market.

However, there is an alternative explanation to this finding. Refugees increase demand in the non-tradable services sector, which could have led to better job openings. Perhaps refugees did not replace natives in regular, salaried jobs: natives *preferred* the non-wage jobs in the services to the regular, salaried jobs in textile. Whereas this story could have been true, I strongly believe that it is inconsistent with the data. First, it is hard to rationalize a

demand shock that leads to only non-salaried employment gains. As the figure shows, there is no increase in wage-employment in Market services. Second, such a demand shock would have drawn natives from many other industries, not solely textile. Yet, I do not see a decrease in formal wage-employment in the other industries. Third, this demand shock cannot explain why both men and women lose their salaried jobs, while it is only the men who make the transition into non-salaried market services. Overall, the evidence does not support the conclusion that natives are leaving their regular, salaried jobs for better opportunities that arise from a demand shock. The evidence strongly suggests that formally employed natives are being displaced by informal refugees in the workforce.

All of the estimates I show in the figures in this section, together with 2SLS estimates using all education-formality-gender-industry-employment type combinations can be found in the Tables E.4, E.5, and E.6 in the Appendix Section E. In general, results are robust across different cuts of the data.

### 5.3 Firm entry

Previous section strongly implies that firms became more informal on the intensive margin by replacing formal native workers with informal refugees as a consequence of the refugee shock. This *informalization* on the intensive margin raises a question about the effects on the extensive margin of informality: whether new entrepreneurs register their businesses. Extensive margin of informality is also an important aspect of most developing countries. It is therefore policy-relevant to study the effect of refugees on firms' decision to be (in)formal.

The identification challenge in this section is more nuanced. First, refugees increase the local population immensely and therefore can increase new firm formation. In contrast, if there are marginal entrepreneurs that are in-between becoming formal or informal, the decrease in informal wages can incentivize these entrepreneurs to remain informal. This would decrease formal firm entry and increase informal firm entry. The empirical challenge is that I do not observe informal firm entry, and hence cannot separately estimate these two effects.

To make progress, I exploit the empirical fact that informal firms are less productive than formal firms (La Porta and Shleifer, 2014; Ulyssea, 2020). Hence, the marginal entrepreneurs should be less productive than non-marginals. Assuming that the demand shock induces new firm formation homogenously across the productivity distribution, there is a testable implication of the informalization effect: we should observe a larger increase in formal entry among productive firms, and a smaller increase, even a decrease in formal entry among small/less productive firms.

To distinguish between more/less productive firms, I first use firms' incorporation status using admin data from Turkstat. Roughly speaking, new firms in Turkey are put into one of the two categories for tax purposes: incorporated firms (tacir) and sole proprietorships (esnaf). The difference between the two types of firms is related to the industry of operation and income. In general, sole proprietorships are smaller in magnitude, and hence more susceptible for informality.

I first estimate equation 4, where the outcome variable is (the natural logarithm of) the number of (i) all firms, (ii) incorporated firms, and (iii) sole proprietorships. Since there are only two periods before treatment I do not control for linear trends. The results are shown in Figure 8a. By 2016, a one standard deviation increase in the instrument is associated with a 7.6% increase in new corporations and no significant change in the number of new sole proprietorships. Since most new firms are sole proprietorships, we do not see an increase in the number of new firms in the aggregate. The 2SLS estimates are shown in columns 1–3 of Figure 8c. A 1 pp increase in refugee to native ratio does not change the number of new firms, increases new corporations by 1.8%, and decreases the number of less productive sole proprietorships by 0.4%. The effect on incorporated firms is statistically significant at the 0.1% level. Taken at face value, these results suggest that refugees caused an increase in the number of new, productive firms; and a mild decrease in the number of new, less productive firms.

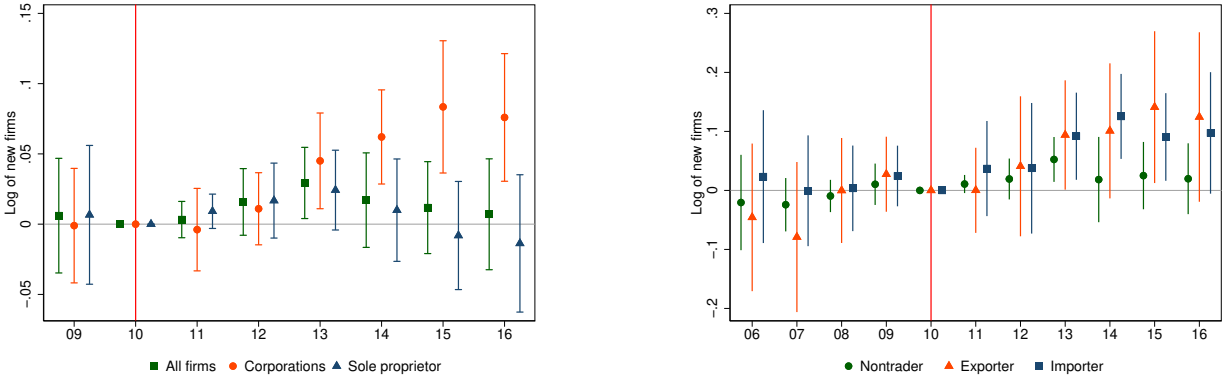
To provide more evidence for this change in the productivity distribution of new firms, I separate firms into three baskets based on their participation in international trade: non-traders, exporters, and importers.<sup>30</sup> The intuition is that firms participating in international trade are more productive than the rest. Hence, the existence of demand and informalization effects would imply that we should observe a higher number of exporter and importer firms, and a smaller, even null effect for non-trader firms. Following the same empirical strategy, I first estimate the reduced-form using equation 4, where the outcome variable is the natural logarithm of the number of (i) non-trader, (ii) exporter, and (iii) importer firms. The results are shown in Figure 8b. Refugees lead to significant increases in the number of both exporter and importer firms, and no change in the number of non-trader firms. The 2SLS estimates are shown in columns 4–6 of Figure 8c. A 1 pp increase in refugee/native ratio causes a 3.2% increase in the number of new, exporter firms, and a 2.0% increase in the number of new, importer firms. It increases the number of non-trader firms only by 0.6%, which is also statistically insignificant.

The null effect of refugees' on non-trader firm entry is even more surprising considering the

---

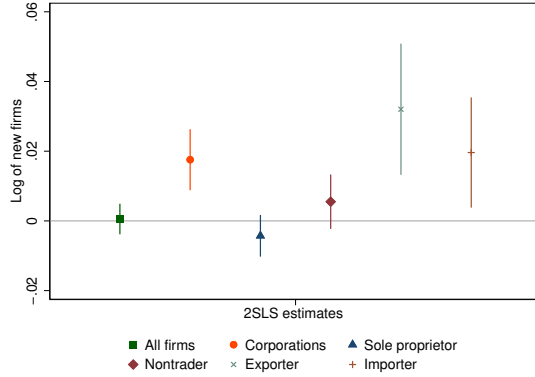
<sup>30</sup>A firm is an exporter (importer) if it appears for at least once in the exports (imports) data during its lifetime.

Figure 8: Effect of refugees on formal firm entry



(a) Tax status

(b) Trade participation



(c) 2SLS estimates

Notes:

effect of population on firm entry, and refugees' increasing the local population substantially. In the Appendix Section D, I first document the strong empirical relationship between new firm formation and population. I then show that refugees substantially increase the total population while causing a very minor decrease in native population (a 1 pp increase in refugee/native ratio decreases the native population by 0.06%).

The heterogeneous effects on the number of new firms across firm types is consistent with a positive effect of immigration on firm entry and an escape to informality among less productive firms. Alternative explanations must rationalize why low-skill immigrants increase the number of only productive firms (such as corporations, or exporter and importer firms), and decrease the number of less productive firms (such as small sole proprietorships).

It is worth emphasizing that without data on informal firms, I cannot credibly conclude that the informal refugee labor supply has incentivized firms to remain informal. However,

to make as much progress as possible without such data, I study refugees' effect on electricity consumption, which is a commonly used indicator to measure informal firm activity (La Porta and Shleifer, 2014). I plot the event-study and IV-DiD estimates in Figure E.14. Controlling for linear trends, I estimate that a 1 pp increase in refugee/native ratio increases the regional electricity consumption by 0.8%. Put differently, I find that whereas refugees did not lead to more firm formation in the aggregate, they caused a sizeable increase in electricity consumption, which would be consistent with more firm activity in the informal sector.

In the Appendix Section G, I provide a tractable model that marries Melitz (2003) and Ulyssea (2018) to formalize my preferred explanation. In the model, heterogeneous firms can exploit two margins of informality: not register their business, and hire workers off the books. Moreover, conditional on registering, firms can also choose to be exporters. The model emphasizes two economic forces that are at play. First, immigrants can induce new firm formation across the productivity distribution via demand and entrepreneurial effects. Second, the informal labor supply shock induces the marginal small firm to remain in the informal sector to obtain easier access to informal workers. These two forces are sufficient to rationalize the reduced form evidence on formal firm entry.

The complexities of this heterogeneous firm model is not necessary to explain the results on the intensive margin. Since the evidence on extensive margin of informality is suggestive, in the model section I focus solely on the results on the intensive margin.

## 6 Model

The purpose of this section is three-folds. First, it relies on a modified version of the canonical labor demand model to determine under what conditions an informal labor supply shock can lead to a decline in formal employment. Second, it quantifies the degree of substitutability between informal and formal workers in the aggregate, an estimate that was missing in the literature on informality. Third, it uses these estimates to analyze a counterfactual in which all refugees were given the right to work. This part relies on assuming the percentage of refugees that can find formal work, a parameter that I cannot estimate in my setting, but it nonetheless provides an answer to an important policy question.

The key finding that will direct the modeling assumptions is that the refugee labor supply shock causes natives to lose both their informal and formal jobs. The former can easily be explained by a downward sloping (informal) labor demand curve (Borjas, 2003). However, the latter requires not only firms' ability to substitute formal workers with informal refugees, but also the displaced (previously informal) natives' limitations in directing their search to

the formal sector. This is why the model will focus on firms' decision to hire formal vs informal workers and abstract away from workers' ability to transition between formal and informal jobs. I discuss the implications for allowing workers' to transition between the two sectors when introduce labor supply.

There is a representative firm that uses informal and formal labor as inputs in production. The hiring costs of formal and informal workers differ due to (1) different wages (e.g., there can be a binding minimum wage for formal workers), and (2) institutional reasons: the firm has to pay a constant payroll tax on formal workers, while it faces an increasing and convex expected cost to hire informal workers, which is summarized by the convex function  $\tau(\cdot)$ . This assumption can be rationalized, for example, by the fact that larger firms have a greater probability of being caught (De Paula and Scheinkman, 2011). This convex cost structure also predicts that the probability of being informally employed should decrease by firm size, which is empirically consistent with the Turkish data. The cost of hiring  $\ell$  formal workers is  $(1 + \tau_w)w_f\ell$ , where  $\tau_w$  is the payroll tax, while the cost of hiring  $\ell$  informal workers is given by  $\tau(\ell)w_i\ell$ .

The firm takes wages as given and produces an homogenous good, whose price is normalized to one.<sup>31</sup> The firm's objective function can be written as follows:

$$\max_{\ell_i, \ell_f} F(\ell_i, \ell_f) - \tau(\ell_i)\ell_i w_i - (1 + \tau_w)w_f\ell_f \quad (6)$$

where  $\tau_w$  is the payroll tax on formal workers, and  $\tau(\ell_i)$  is the expected cost of hiring informal workers. In particular, I assume that  $\tau(\ell_i) = \ell_i^\gamma$  with  $\gamma > 0$ , which satisfies the convex cost structure assumed in the literature (Ulyssea, 2018). The production function  $F$  has a CES form.

$$F(\ell_i, \ell_f) = (\eta\ell_i^\rho + (1 - \eta)\ell_f^\rho)^{\frac{\alpha}{\rho}}$$

where  $0 < \alpha < 1$  indicates a decreasing returns to scale (in labor) production function, that is appropriate to study short-run adjustments;  $\sigma = \frac{1}{1-\rho}$  is the elasticity of substitution between formal and informal labor, and  $\eta$  is the share parameter of informal labor input. Unlike the recent structural work that has to assume the elasticity of substitution between the formal and informal workers, the cost shock to the informal workers through refugee labor supply allows me to estimate this parameter.

---

<sup>31</sup>The competitive market assumption simplifies the algebra, but can be opposed due to the various frictions in the labor markets of developing economies. The implications of monopsony, and how it can interact with informality is beyond the scope of this paper.

Given this set up, the first order conditions of a profit maximizing firm is given by:

$$\begin{aligned}\alpha\eta\ell_i^{\rho-1-\gamma}Y^{\frac{\alpha-\rho}{\alpha}} &= w_i(1+\gamma) \\ \alpha(1-\eta)\ell_f^{\rho-1}Y^{\frac{\alpha-\rho}{\alpha}} &= w_f(1+\tau_w)\end{aligned}\tag{7}$$

where  $Y = (\eta\ell_i^\rho + (1-\eta)\ell_f^\rho)^{\frac{\alpha}{\rho}}$  is the output produced by the firm. Given wages  $w_i$  and  $w_f$ , the labor demand for informal workers,  $L_i^d(w_i, w_f)$ , and formal workers,  $L_f^d(w_i, w_f)$ , are given by equation 7.

## 6.1 Equilibrium

To close the model, I need to specify the labor supply. Let  $L_i^{N,S}(w_i)$  and  $L_f^{N,S}(w_f)$  be the informal and formal labor supply curves of natives. Notice that labor supply curves in either sector are independent of the wages in the other sector. This is a simplifying assumption that eliminates workers' ability to search for both informal and formal jobs. Allowing for workers' to direct their search endogenously would reduce the effective increase in informal labor supply due to the refugee shock, and therefore would limit the adjustments in the labor demand that the model isolates.<sup>32</sup>

In equilibrium, labor markets must clear: informal and formal wages are such that labor supply equals labor demand.

$$\begin{aligned}L_i^S(w_i) &= L_i^D(w_i, w_f) \\ L_f^S(w_f) &= L_f^D(w_i, w_f)\end{aligned}\tag{8}$$

## 6.2 The effect of an informal labor supply shock

In this competitive model, the effect of an informal labor supply shock on labor demand can be captured by the elasticities of informal and formal labor demand w.r.t. informal wages (assuming formal wages are fixed by a minimum wage for simplicity). After some algebra, one can show that the elasticities of informal and formal labor demand w.r.t. informal wages are given by:

$$\begin{aligned}\epsilon_{L_i, w_i} &= -\frac{1-\rho-(\alpha-\rho)s_f}{(1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f+(1-\rho)s_i]} \\ \epsilon_{L_f, w_i} &= -\frac{(\alpha-\rho)s_i}{(1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f+(1-\rho)s_i]}\end{aligned}\tag{9}$$

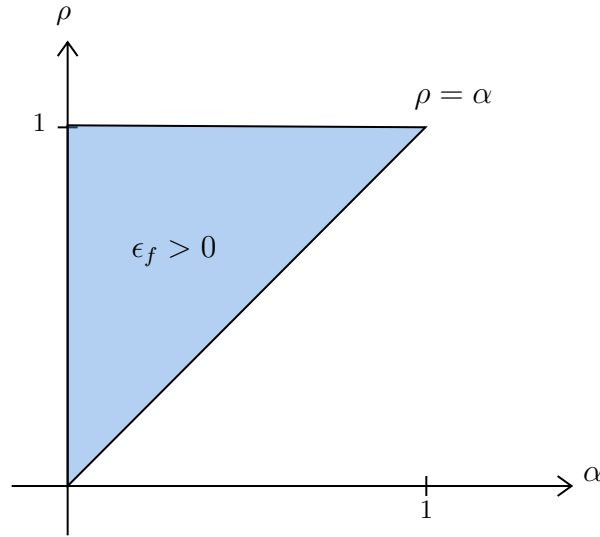
---

<sup>32</sup>The interested reader can read Meghir et al. (2015) for a search model in which workers can search for jobs in both the formal and informal sectors.

where  $s_i = \frac{\eta L_i^\rho}{\eta L_i^\rho + (1-\eta)L_f^\rho}$  is the informal share in the production, and vice versa for  $s_f$ .

Equation 9 formalizes two intuitive results. First,  $\epsilon_{L_i, w_i} < 0$  for all potential values of  $\rho$ , meaning as informal wages decrease, firms demand more informal labor. However, the effect on the formal labor demand is more nuanced. The sign of the elasticity of formal labor demand w.r.t. informal wages (which I will denote shortly as  $\epsilon_f$ ) depends solely on the sign on  $\alpha - \rho$ . This relationship is plotted on Figure 9. When the labor share of production  $\alpha$  is less than  $\rho$ , the elasticity of formal labor demand is positive. This means that when informal wages go down in the economy, formal labor demand also goes down.

Figure 9: Set of parameters by which the elasticity of formal labor demand w.r.t. informal wages is negative



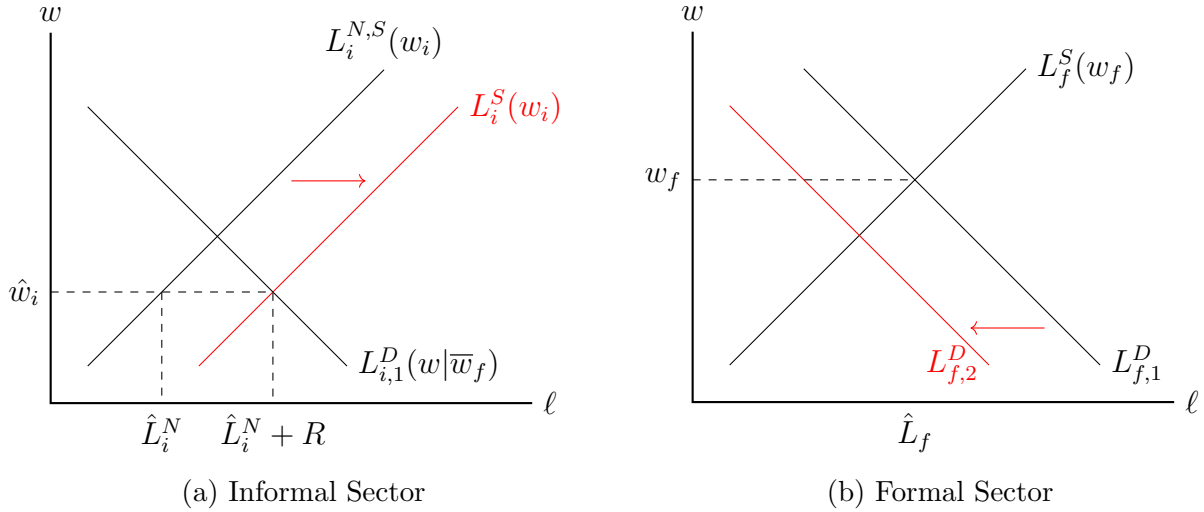
To understand the intuition behind this result, consider the change in the marginal productivity of a formal worker when an informal worker is hired:

$$\frac{\partial(\log \frac{\partial F}{\partial L_f})}{\partial L_i} = (\alpha - \rho)L_i s_i$$

In the case of a CRTS production function ( $\alpha = 1$ ) and formal and informal workers not being perfect substitutes ( $\rho < 1$ ), hiring an informal worker makes formal workers *more* productive due to the complementarity between workers. Consequently, for the same formal wage, the firm demands more formal labor, giving us a negative elasticity of formal labor demand  $\epsilon_{L_f, w_i} < 0$ . However, as  $\alpha$  decreases, hiring an additional worker incurs productivity losses for the rest of the workers due to decreasing returns. If  $\alpha$  is small enough (i.e.,  $\alpha < \rho$ ), then the productivity loss from technological constraints (e.g., capital being constant in the short run) overpowers the productivity gain from the complementarity between workers.



Figure 10: Equilibrium with informal labor supply shock



Notes: For illustrative purposes, I assume the formal wage is fixed by a binding minimum wage. Otherwise, a decrease in the effective formal wage would also push the informal labor demand curve slightly upwards.

Consequently, an informal labor supply shock that reduces informal wages can incentivize firms to substitute away from formal workers.<sup>33</sup>

Figure 10 shows how the refugee labor supply impacts the labor market equilibrium in this model where the production function has decreasing returns to scale and formal and informal labor are highly substitutable (i.e.,  $\alpha < \rho$ ). The left panel shows the equilibrium for informal workers and the right panel shows the equilibrium for formal workers. For ease of exposition, I assumed that refugees supply labor inelastically. This results in a parallel shift in the informal labor supply curve, causing (1) a decline in informal wages, (2) a decline in native informal employment, and (3) an increase in the aggregate informal employment. Since formal and informal workers are substitutes, the decrease in the informal wages incentivizes formal firms to rely more intensively on informal workers. This shifts the formal labor demand curve inward. As firms reduce their demand for formal workers, the amount of native formal employment decreases, despite refugees being unable to work formally. This explains why the informal labor supply shock of Syrian refugees has led to decreases in both informal and formal markets.

<sup>33</sup>An alternative way to generate this qualitative prediction is presented in Delgado-Prieto (2021), who incorporates a CRTS (in labor) production function with imperfect competition in that the price is determined by product demand into a framework similar to Ulyssea (2018). In his model, an increase in the number of informal workers can reduce the productivity of existing employees by lowering the price. This is different from the approach here. My model achieves the same results through a different mechanism, and moreover it does so in a simpler fashion and without introducing an additional free parameter (the elasticity of product demand).

### 6.3 Estimating the elasticity of substitution between informal and formal labor

Previous section shows under what conditions an informal labor supply shock can cause disemployment in the formal sector. In this section, I map some of the empirical estimates into model parameters. Let  $\kappa$  denote the relative elasticities of informal and formal labor demand:  $\kappa \equiv -\frac{\epsilon_{L_i, w_i}}{\epsilon_{L_f, w_i}}$ . Notice that this ratio is informative of the elasticity of the substitution between the informal and formal workers:

$$\begin{aligned}\kappa &\equiv -\frac{\epsilon_{L_i, w_i}}{\epsilon_{L_f, w_i}} = \frac{1 - \rho + (\rho - \alpha)(1 - s_i)}{(\rho - \alpha)s_i} \\ \implies \rho &= \alpha + \frac{1 - \alpha}{s_i(1 + \kappa)}\end{aligned}\tag{10}$$

Hence, an estimate of the labor share in production  $\alpha$ , informal labor share in production  $s_i$ , and the ratio of elasticities  $\kappa$  allow me to back up an estimate of  $\rho$ .<sup>34</sup>

Coming back to the thought experiment in section 5.1, consider a region with 1000 working-age low-skill natives. On average, 23.3% of low-skill natives are wage-employed, with 8.5% being informal and 14.8% being formal. This gives 85 informally and 148 formally employed workers. Let this region receive 100 refugees, a 10 pp increase in refugee/native ratio. Survey evidence suggests that 45 refugees end up working informally in wage-jobs. My estimates suggest that this shock leads to 23 natives losing informal jobs, and 13 natives losing formal jobs. This leads to 107 informally employed, and 135 formally employed workers in the economy. As a consequence of the informal labor supply shock, the number of informal workers in the economy has increased by 25.9%, and formal employment has decreased by 8.8%. Consequently, I estimate that the ratio of elasticities is:  $\hat{\kappa} = \frac{25.9}{8.8} = 2.95$

I further calibrate  $\alpha = 2/3$ , the canonical labor share in production, and  $\eta = 0.5$ .<sup>35</sup> Then, I solve for  $\rho$  using equation 10. I find that  $\hat{\rho} = 0.89$ , which implies an elasticity of substitution between informal and formal labor of  $\sigma_{i,f} \approx 9$ . This means that informal and formal workers are highly substitutable in the aggregate economy. This result is intuitive also at the micro level. Consider a firm having 10 workers, 3 informal and 7 formally employed. The firm hires an 11th worker to do certain tasks. Whether the firm hires that person formally or informally does not change the tasks performed by the worker. The only set of parameters that would allow for the worker to have the same productivity regardless of formality status

---

<sup>34</sup>Recall that  $s_i = \frac{\eta L_i^\rho}{\eta L_i^\rho + (1-\eta)L_f^\rho}$ , and hence is decreasing in  $\rho$  so long as  $L_f > L_i$  (which is empirically the case). Therefore, for any estimate of  $\eta$ , there exist a unique  $\rho$  for which equation 10 holds.

<sup>35</sup>This can be justified, for instance, if informal and formal workers were perfect substitutes, and if formalizing a worker would not change his/her productivity at the job. In that setting, hiring a worker informally or formally should lead to the same increase in production, which is only possible if  $\eta = 0.5$ .

is  $\rho = 1$  (i.e., perfect substitutes), and  $\eta = 0.5$ , which is already calibrated.

## 6.4 Estimating Labor demand elasticities and the change in informal wages

Given the estimated and calibrated parameters, only the parameter controlling the convex cost of hiring informal workers,  $\gamma$ , is needed to estimate the labor demand elasticities in equation 9. Following the modeling assumptions in *Ulysses* (2018), I assume that  $\gamma = 1$ . This implies that the elasticity of informal labor demand w.r.t. informal wages,  $\epsilon_{L_i, w_i}$ , is -0.87, and the elasticity of formal labor demand w.r.t. formal wages,  $\epsilon_{L_f, w_i}$ , is 0.29.

The labor demand elasticities also allow me to back out the decrease in informal wages faced by firms. Given enough data, I could estimate how much informal wages have gone down due to the refugee shock directly from the data. However, two problems prevent me from performing this reduced-form analysis. First, I do not observe the wages that refugees work for in the data. I only know refugees' average monthly earnings in 2019. For instance, refugees and natives could be working for different wages in practice, reducing the labor costs of firms. This would not appear in the Turkish labor force surveys as refugees are not sampled from. Second, I document sizeable decreases in employment rates among natives, which means the distribution of wages in the economy might change. If the employment losses occurred mostly among the low earners, then the compositional effect could increase the average native wages observed in the data while the wages faced by firms are decreasing. Consequently, there is no reduced-form way of estimating the decrease in wages faced by firms given the available data.

Instead, I rely on the model to back out the change in wages. I can approximate the increase in total informal wage-employment in the economy by  $\Delta \ln L_i \approx \epsilon_{L_i, w_i} \times \Delta \ln w_i$ . As a consequence of the refugee shock, I estimated that the total informal employment in the economy has increased by 25.9%. Together with the estimated informal labor demand elasticity of  $-0.87$ , the model suggests that informal wages faced by firms has decreased by 30%. Interestingly, this is very similar to the relative monthly earnings of refugees and natives in 2019. I estimate that informally employed natives made on average 1565 TRY per month; and Turkish Red Crescent and WFP (2019) find that refugees on average made 1058 TRY per month, 32% less than comparable natives.

## 6.5 Counterfactual: What would have happened if refugees had work permits?

The empirical evidence that I put forward in this paper, the effect of informal workers on formal sector, naturally raises an important policy question: what would have happened if refugees were given the right to work formally? Would more natives lose jobs as a consequence of this policy, and would there be less incentives for firms to be informal? To answer these questions, I need to predict the percentage of refugees that would work formally if provided with work permits. For instance, all natives in Turkey have the right to find formal work, but only 64% of low-skill natives actually work formally. Naturally, if all refugees were given work permits, not all would find formal work. The ratio of refugees that can work formally is the key parameter that determines the bite of providing work permits to refugees.

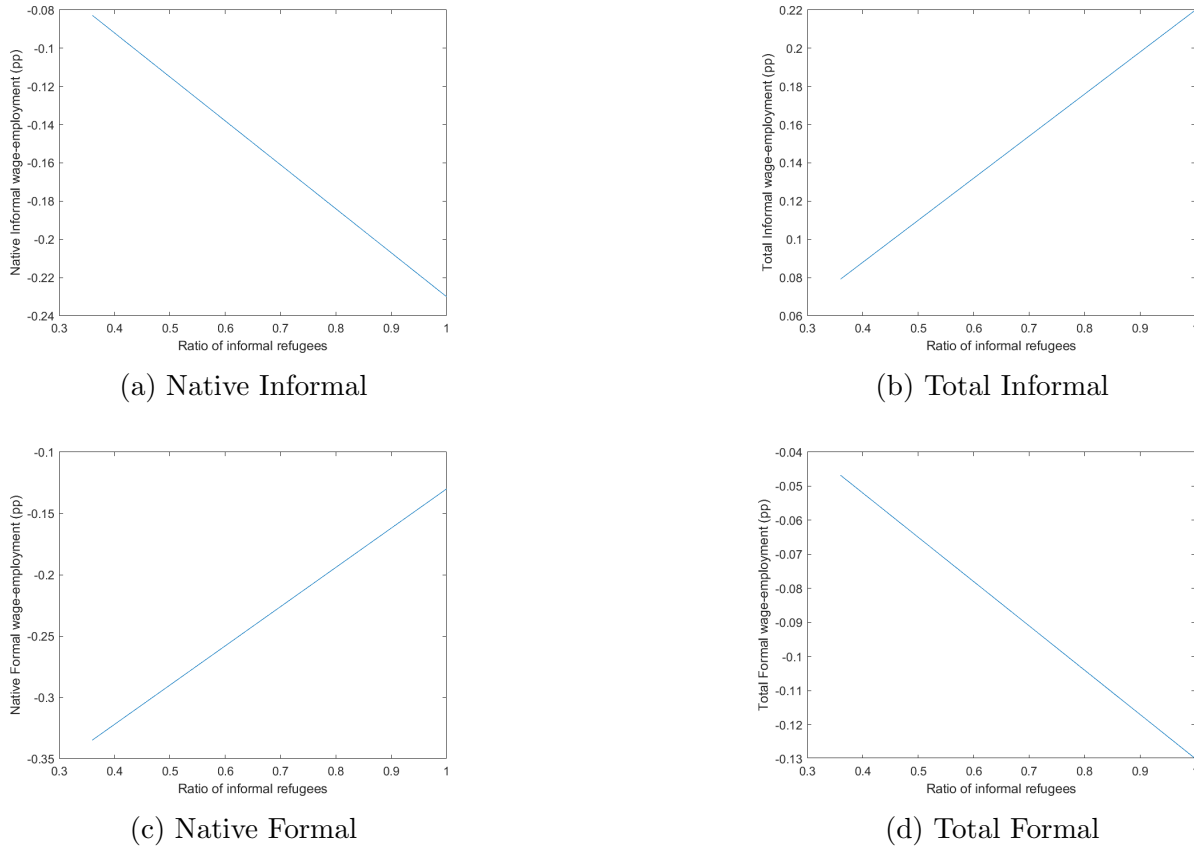
More formally, let  $c \in [0, 1]$  denote the ratio of refugees that are endowed with only informal labor. That is, if  $c = 1$ , all refugees can only work informally, and if  $c = 0$ , all refugees can work formally. The key observation here is that  $c$  determines the relative shifts in the informal and formal labor supply curves. As  $c$  starts decreasing, the informal labor supply curve shifts right less, the informal wages decrease less, the conditional formal labor demand curve shifts down less, and the formal labor supply curve shifts right more. The effect of these movements on native (and total) informal (and formal) employment can be approximated by  $c$  and the already estimated effects from the empirical section.

Consider a 1 pp increase in refugee/native ratio where  $100 \cdot c\%$  of refugees are endowed with only informal labor. Using the estimates from the previous sections and relying on linear approximations, I estimate the effect of a 1 pp increase in refugee/native ratio on native and total employment outcomes as:

$$\begin{aligned} \text{Native informal employment (pp): } & (-0.23) \times c \\ \text{Total informal employment (pp): } & (-0.23 + 0.45) \times c \\ \text{Native formal employment (pp): } & (-0.13 \times c) - 0.45 \times (1 - c) \\ \text{Total formal employment (pp): } & (-0.13) \times c \end{aligned} \tag{11}$$

Consequently, given  $c$ , one can estimate the counterfactual of what would have happened if all refugees were given work permits. Unfortunately, there is not a good data-driven way to estimate  $c$ . In Turkey, there are very few refugees who were provided with work permits (about whom we don't have data), and they are selected. Therefore, I cannot credibly estimate the underlying formality levels of refugees from the data. Instead, I assume that the underlying level of informality of refugees is weakly higher than that of the natives (i.e.,

Figure 11: Effects of a 1 pp increase in refugee/native ratio with different levels of refugee informality



$c \in [0.36, 1]$ , which is a conservative assumption), and I show the counter-factual effects of providing *all* refugees with work permits for all potential values of  $c$  in Figure 11. As a benchmark, if refugees had the same informality levels as the natives, I predict that a 1 pp increase in refugee/native ratio would have caused a 0.08 pp decrease in native informal employment, 0.08 pp increase in total informal employment, 0.33 pp decrease in native formal employment, and only a 0.05 pp decrease in total formal employment (as opposed to the 0.13 pp decrease I estimated in the reduced-form). Intuitively, as more refugees can find formal jobs, fewer natives lose informal jobs, and more natives lose formal jobs.

Put differently, these estimates suggest that not providing work permits to refugees cost tax revenue to the host countries through reduced formal employment. For example, in 2011 there were 50 million natives in Turkey between the ages of 15-65. 67.5% of them (i.e., 33.75 million) were not in school and had less than a high-school degree. By 2016, refugees had increased Turkey's overall population by 4 pp. Using the estimates on Figure 11d and the benchmark case of refugees having the same informality level of low-skill Turkish natives,

I conclude that not providing work permits to refugees have caused approximately 112,000 jobs formal jobs to disappear. At the time, the formal monthly minimum wage was around \$549 before tax and \$433 after tax. Assuming that all the jobs lost were minimum-wage jobs (as I did in the model), not providing work permits to refugees has caused 155 million USD in personal income tax revenue to Turkey in 2016 only. In reality, there were likely more informalizing incentives that affected tax revenue that I cannot capture in my setting (e.g., firms' choosing to remain smaller to avoid detection while hiring informal workers). Future work can shed more light to the extent of the frictions created by the lack of work permits.

## 7 Conclusion

The number of refugees has tripled in the last decade, raising a lot of policy challenges and variation to study. One contentious aspect has been the formal labor market access of refugees. Soon after the Russian invasion of Ukraine began, Poland declared that Ukrainian refugees in Poland would have work permits (Lesinska, 2022). Colombia gave work permits to the majority of Venezuelan refugees 4 years after the crisis began (Bahar et al., 2021). Most Syrian refugees in Turkey do not have work permits to this day. Despite the importance and the prevalence of this policy variation, not much has been written about or using the issue of work permits in Economics.

In this paper, I use Syrian refugees' informality both to study the dynamics between informal and formal sectors and to quantify the aggregate effect of work permits on the labor markets. By employing an IV-DiD strategy where I instrument the refugee location choice by a distance-based instrument, I show that an informal labor supply shock reduces both the informal and formal wage-employment rates of the natives. Additionally, I show that this disemployment comes solely from the industries that refugees work at, which highlights the labor supply mechanism as the main channel behind my results. On formal firm entry, I show that refugees increase the number of new, exporter and importer firms, do not change the number of new, non-trader firms, and weakly decrease the number of sole proprietorships (esnaf) that are small firms with limited income. Put differently, refugees lead to a change in the productivity distribution of new formal firms. I show that this change is consistent with small marginal firms choosing to remain informal, but the lack of data on informal firms prevents me from testing this directly.

To explain these findings, I propose a simple extension of the canonical labor demand framework in which the representative firm uses a CES aggregate of informal and formal labor in production. I first show under what conditions an informal labor supply can lead to a decrease in formal employment. Then, I use the reduced-form estimates to estimate the

elasticity of substitution between informal and formal labor in the aggregate, which I find to be around 9. This implies that informal and formal labor are highly substitutable in the aggregate. To the best of my knowledge, this is the first paper that estimates this elasticity using an IV-DiD design. Lastly, I use this simple model to quantify the role of work permits in refugee crises. Under a benchmark case in which refugees have the same informality level as of the natives, I estimate that providing work permits to refugees would have created 112,000 formal jobs and increased government tax revenue by \$155 million per year.

From an economics perspective, my results imply that firms' ability to substitute between formal and informal workers is *more powerful* than workers' ability to transition between formal and informal jobs in Turkey. The latter would have caused the displaced, previously informal natives to transition to formal jobs, which would have increased native formal employment rate. Instead, I document native disemployment, which implies that formal firms substitute formal natives with informal refugees. Future work around formalization efforts should take firms' ability to substitute between formal and informal workers into account.

From a policy perspective, my results emphasize an equilibrium force that is of first order importance, yet absent in the policy discussions around the labor market access of refugees (Clemens et al., 2018). Informal refugees incentivizing firms to be more informal both on the intensive and extensive margins can exacerbate the problems related to having an informal sector. However, providing work permits may not be sufficient to oppose these economic forces. For instance, Colombia gave work permits to the Venezuelan refugees in 2018, but this did not lead to major changes in the formal employment rates of natives, at least in the short term (Bahar et al., 2021). There is still a lot that we don't know about the role of work permits in refugee crises. Future research should shed more light on the general equilibrium effects of work permits, and help determine the optimal policy for host countries.

## References

- Afanasieva, Dasha, “Turkey will not give Syrian refugees right to work: Minister,” <https://www.hurriyetdailynews.com/turkey-will-not-give-syrian-refugees-right-to-work-minister-86642> 2015. Accessed: 2022-07-01.
- Akgündüz, Yusuf Emre and Huzeyfe Torun, “Two and a half million Syrian refugees, tasks and capital intensity,” *Journal of Development Economics*, 2020, 145, 102470.
- Akgunduz, Yusuf Emre, Yusuf Kenan Bayir, Seyit Mumin Cilasun, and Murat Guray Kirdar, “Consequences of a Massive Refugee Influx on Firm Performance and Market Structure,” Technical Report, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London 2022.
- Aksu, Ege, Refik Erzan, and Murat Güray Kirdar, “The impact of mass migration of Syrians on the Turkish labor market,” *Labour Economics*, 2022, p. 102183.
- Altındağ, Onur, Ozan Bakış, and Sandra V Rozo, “Blessing or burden? Impacts of refugees on businesses and the informal economy,” *Journal of Development Economics*, 2020, 146, 102490.
- Amaral, Pedro S and Erwan Quintin, “A competitive model of the informal sector,” *Journal of monetary Economics*, 2006, 53 (7), 1541–1553.
- Andrade, Gustavo Henrique De, Miriam Bruhn, and David McKenzie, “A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms,” *The World Bank Economic Review*, 2016, 30 (1), 24–54.
- Angrist, Joshua D and Adriana D Kugler, “Protective or counter-productive? labour market institutions and the effect of immigration on natives,” *The Economic Journal*, 2003, 113 (488), F302–F331.
- Bahar, Dany, Ana María Ibáñez, and Sandra V Rozo, “Give me your tired and your poor: Impact of a large-scale amnesty program for undocumented refugees,” *Journal of Development Economics*, 2021, 151, 102652.
- Borjas, George J, “The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market,” *The quarterly journal of economics*, 2003, 118 (4), 1335–1374.



- , “Immigration economics,” in “Immigration Economics,” Harvard University Press, 2014.
- **and Joan Monras**, “The labour market consequences of refugee supply shocks,” *Economic Policy*, 2017, *32* (91), 361–413.
- Bosch, Mariano and Julen Esteban-Pretel**, “Job creation and job destruction in the presence of informal markets,” *Journal of Development Economics*, 2012, *98* (2), 270–286.
- Card, David**, “The impact of the Mariel boatlift on the Miami labor market,” *ILR Review*, 1990, *43* (2), 245–257.
- Carpio, Ximena V Del and Mathis Wagner**, *The impact of Syrians refugees on the Turkish labor market*, The World Bank, 2015.
- Cengiz, Doruk and Hasan Tekgüç**, “Is it merely a labor supply shock? Impacts of Syrian migrants on local economies in Turkey,” *ILR Review*, 2022, *75* (3), 741–768.
- Ceritoglu, Evren, H Burcu Gurcihan Yunculer, Huzeyfe Torun, and Semih Tumen**, “The impact of Syrian refugees on natives’ labor market outcomes in Turkey: evidence from a quasi-experimental design,” *IZA Journal of Labor Policy*, 2017, *6* (1), 1–28.
- Clemens, Michael, Cindy Huang, Jimmy Graham et al.**, “The economic and fiscal effects of granting refugees formal labor market access,” *Center for Global Development Working Paper*, 2018, *496*.
- Delgado-Prieto, Lukas**, “Dynamics of Local Wages and Employment: Evidence from the Venezuelan Immigration in Colombia,” 2021.
- Demirci, Murat and Murat Güray Kırdar**, “The labor market integration of Syrian refugees in Turkey,” *World Development*, 2023, *162*, 106138.
- Erten, Bilge and Pinar Keskin**, “Female employment and intimate partner violence: evidence from Syrian refugee inflows to Turkey,” *Journal of Development Economics*, 2021, *150*, 102607.
- Friedberg, Rachel M**, “The impact of mass migration on the Israeli labor market,” *The Quarterly Journal of Economics*, 2001, *116* (4), 1373–1408.
- Hunt, Jennifer**, “The impact of the 1962 repatriates from Algeria on the French labor market,” *ILR Review*, 1992, *45* (3), 556–572.
- İçduygu, Ahmet**, “Syrian refugees in Turkey: The long road ahead,” *Washington, DC: Migration Policy Institute*, 2015.

- Lebow, Jeremy**, “The labor market effects of Venezuelan migration to Colombia: reconciling conflicting results,” *IZA Journal of Development and Migration*, 2022, 13 (1).
- Lesinska, Magdalena**, “Poland: Parliament adopts law on assistance to Ukrainian refugees,” [https://ec.europa.eu/migrant-integration/news/poland-parliament-adopts-law-assistance-ukrainian-refugees\\_en](https://ec.europa.eu/migrant-integration/news/poland-parliament-adopts-law-assistance-ukrainian-refugees_en) 2022. Accessed: 2022-06-21.
- Meghir, Costas, Renata Narita, and Jean-Marc Robin**, “Wages and informality in developing countries,” *American Economic Review*, 2015, 105 (4), 1509–46.
- Melitz, Marc J**, “The impact of trade on intra-industry reallocations and aggregate industry productivity,” *econometrica*, 2003, 71 (6), 1695–1725.
- Monteiro, Joana CM and Juliano J Assunção**, “Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises,” *Journal of Development economics*, 2012, 99 (1), 105–115.
- Ozar, Samsa**, “Micro and small enterprises in Turkey: Uneasy development,” in “Research report series/Economic Research Forum; no. 0420” Economic Research Forum, Cairo, EG 2003.
- Paula, Aureo De and José A Scheinkman**, “The informal sector: An equilibrium model and some empirical evidence from Brazil,” *Review of Income and Wealth*, 2011, 57, S8–S26.
- Peri, Giovanni and Vasil Yassenov**, “The labor market effects of a refugee wave synthetic control method meets the mariel boatlift,” *Journal of Human Resources*, 2019, 54 (2), 267–309.
- Porta, Rafael La and Andrei Shleifer**, “Informality and development,” *Journal of economic perspectives*, 2014, 28 (3), 109–26.
- Rauch, James E**, “Modelling the informal sector formally,” *Journal of development Economics*, 1991, 35 (1), 33–47.
- Rocha, Rudi, Gabriel Ulyssea, and Laísa Rachter**, “Do lower taxes reduce informality? Evidence from Brazil,” *Journal of development economics*, 2018, 134, 28–49.
- Schramm, Heidi R**, “The equilibrium effects of income taxation on formal and informal labor markets,” in “NEUDC 2014 Working Paper. Presented at the Northeast Universities Development Consortium Conference” Boston University Boston 2014.

- Tumen, Semih**, “The economic impact of Syrian refugees on host countries: Quasi-experimental evidence from Turkey,” *American Economic Review*, 2016, *106* (5), 456–60.
- Turkish Red Crescent and WFP**, “Refugees In Turkey: Livelihoods Survey Findings,” 2019. <https://reliefweb.int/report/turkey/refugees-turkey-livelihoods-survey-findings-2019-entr>.
- Ulyssea, Gabriel**, “Firms, informality, and development: Theory and evidence from Brazil,” *American Economic Review*, 2018, *108* (8), 2015–47.
- , “Informality: Causes and consequences for development,” *Annual Review of Economics*, 2020, *12*, 525–546.
- UNHCR**, “UNHCR refugee statistics,” <https://www.unhcr.org/refugee-statistics/> 2021. Accessed: 2022-06-15.
- , “Operational Data Portal Refugee Situations,” 2022. <https://data2.unhcr.org/en/situations/syria/location/113> [Accessed: 2022-10-24].
- World Bank**, *Doing business 2011: Making a difference for entrepreneurs*, The World Bank, 2010.

# A Data

Table A.1: HLFS Summary Statistics

Formality	Wage Employment				Non-wage Employment			
	All	Informal	Formal		All	Informal	Formal	
Skill			Low	High			Low	High
Panel A: Aggregate								
Pooled	0.323	0.071	0.157	0.459	0.188	0.124	0.061	0.071
Men	0.491	0.106	0.292	0.544	0.251	0.134	0.122	0.107
Women	0.160	0.037	0.045	0.340	0.127	0.115	0.010	0.020
Panel B: Across industries								
Agriculture	0.011	0.009	0.002	0.002	0.101	0.085	0.021	0.006
Textile	0.028	0.008	0.021	0.018	0.006	0.004	0.002	0.002
Other manufacturing	0.062	0.008	0.042	0.081	0.010	0.004	0.006	0.007
Construction	0.028	0.012	0.016	0.015	0.007	0.003	0.003	0.005
Market Services	0.110	0.023	0.056	0.155	0.056	0.024	0.026	0.047
Non-market Services	0.084	0.011	0.020	0.188	0.008	0.004	0.004	0.005

*Note:* Household Labor Force Surveys between 2004–2016 are used. Wage employment is defined as regular, salaried work. Non-wage employment is defined as the rest, which is the combination of self-employment, unpaid family work, and being an employer. Skill levels are determined by education. Low-skill refers to people without high-school degrees. High-skill refers to people with at least high-school degrees. Industry specifications follow the ISIC categories.

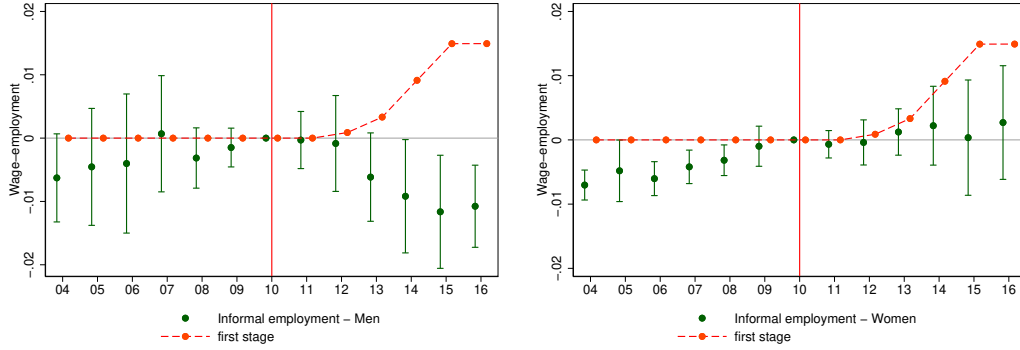
Table A.2: The weights assigned to Syrian regions

Governorate	Pop share	Share in Turkey	IV
Aleppo	24.2	35.7	42.5
Idleb	8.5	20.9	15.4
Raqqa	4.1	10.9	5.8
Lattakia	5.0	9.2	7.7
Hama	8.6	7.5	5.9
Hassakeh	6.5	5.4	9.3
Dayr Az Zor	6.9	3.9	4.8
Damascus	15.0	3.8	2.7
Homs	8.7	1.7	2.8
As Suweida	2.0	0.4	0.4
Daraa	4.6	0.3	0.2
Al Qunaytirah	2.0	0.1	0.4
Tartous	3.9	0.1	2.0

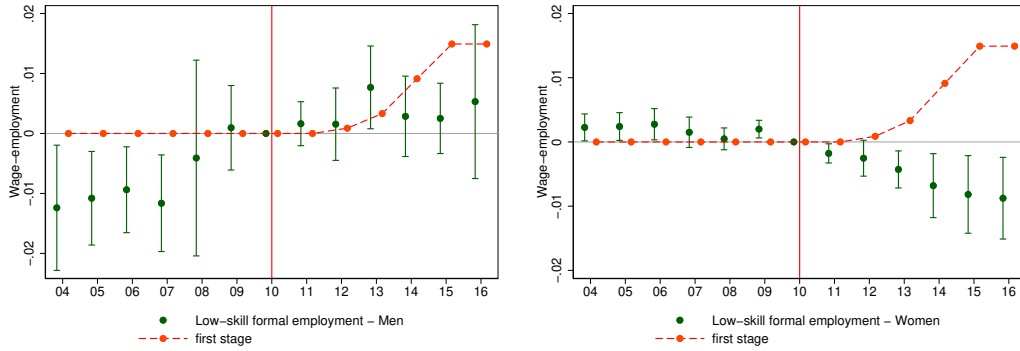
## B Additional Event-study figures

Figure B.1: Event-study figures with raw data

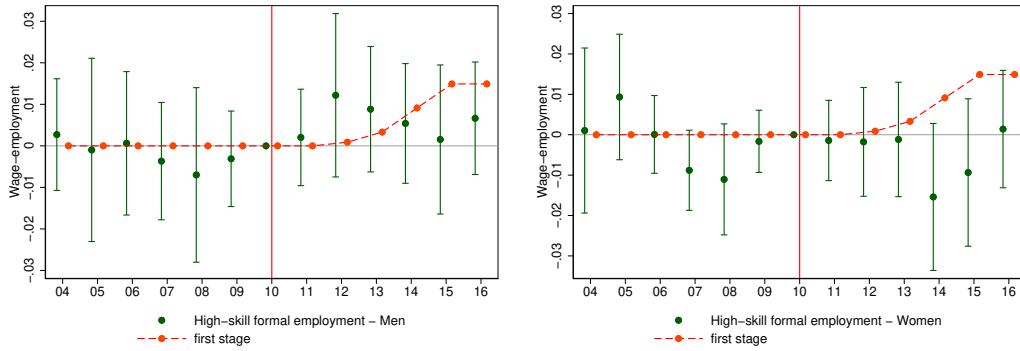
(a) Informal wage-employment



(b) Formal low-skill wage-employment

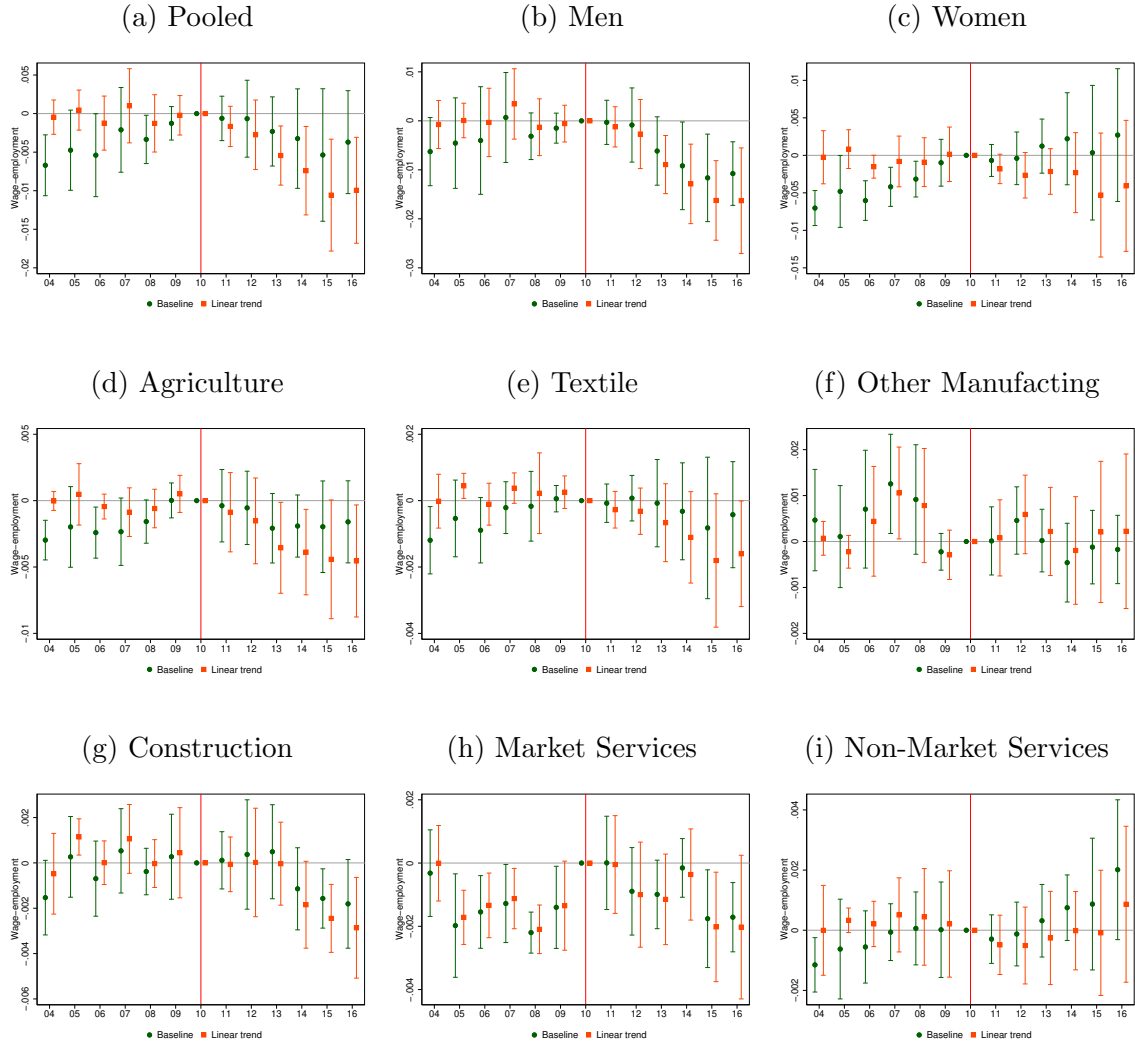


(c) Formal high-skill wage-employment



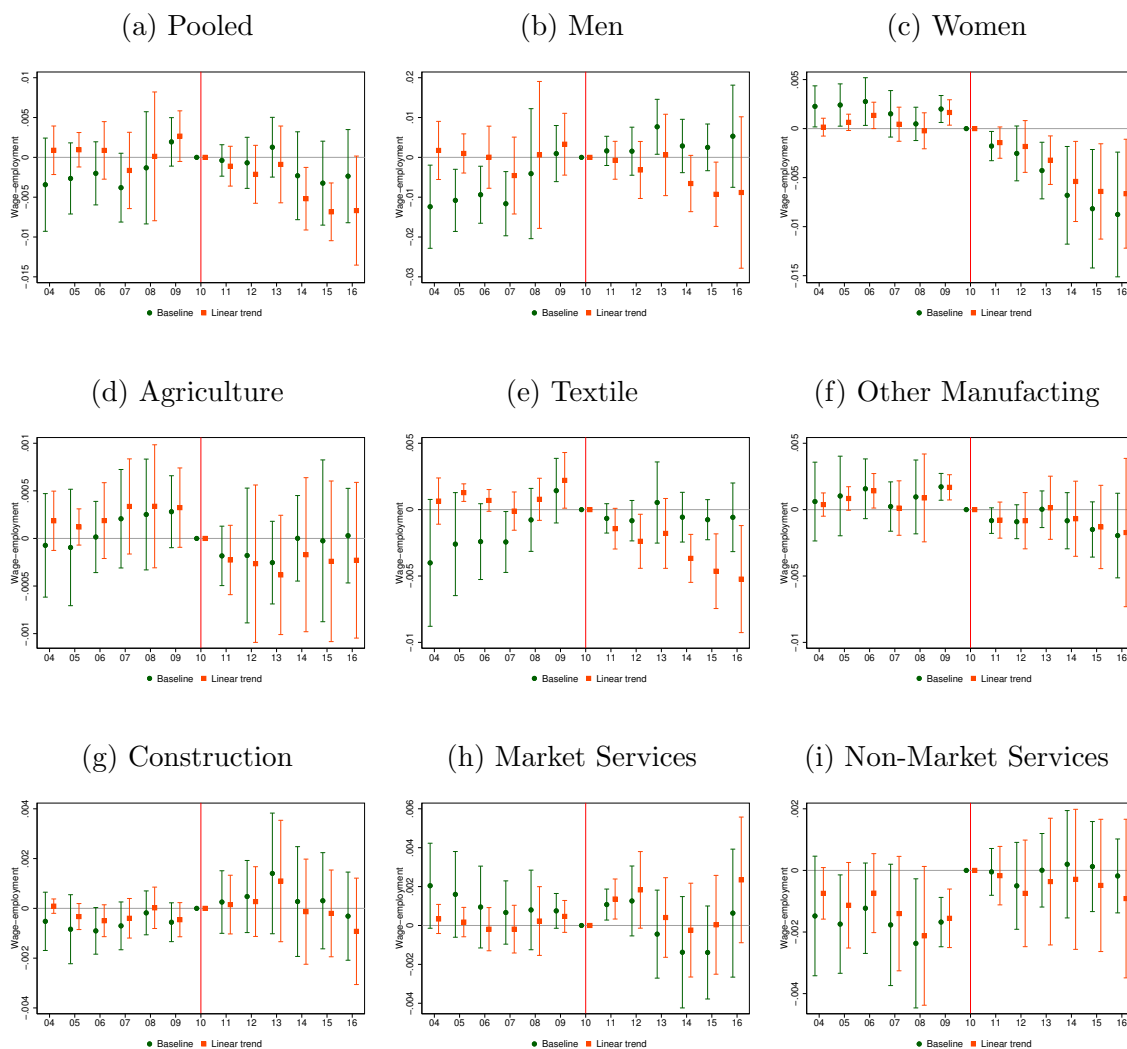
Notes: The reduced-form estimates come from the regression  $y_{p,t} = \sum_{j \neq 2009} \beta_j (\text{year}_j \times \tilde{Z}_p) + \delta_p + \delta_t + \epsilon_{p,t}$ , where  $\tilde{Z}_p$  is the standardized version of the instrument. The outcome variable is the informal wage-employment rate in Panel A, formal wage-employment rate for the low-skilled in Panel B, and formal wage-employment rate for the high-skilled in Panel C. The left column shows the estimates for men, and the right column for women. Standard errors are clustered at nuts2 region level. The 95% confidence interval is plotted. The red dashed line shows the event-study point estimates for the first-stage, which captures how the treatment intensity predicted by the instrument varies over time.

Figure B.2: Event-study figures for informal wage employment estimates in Figure 5



Notes: The reduced-form estimates come from the regression  $y_{p,t} = \sum_{j \neq 2009} \beta_j (\text{year}_j \times \tilde{Z}_p) + \delta_p + \delta_t^s + \epsilon_{p,t}$ , where  $\tilde{Z}_p$  is the standardized version of the instrument. In the baseline specification, the outcome variable is the raw wage employment rate. In the “linear trend” specification, I use the detrended outcome, where I estimate the trend in a regression using data between 2004-2010:  $y_{p,t} = f_p + f_t + y_{tr,p} * t + \xi_{p,t}$ . The industry specifications follow the ISIC standards. Standard errors are clustered at nuts2 region level. The 95% confidence interval is plotted.

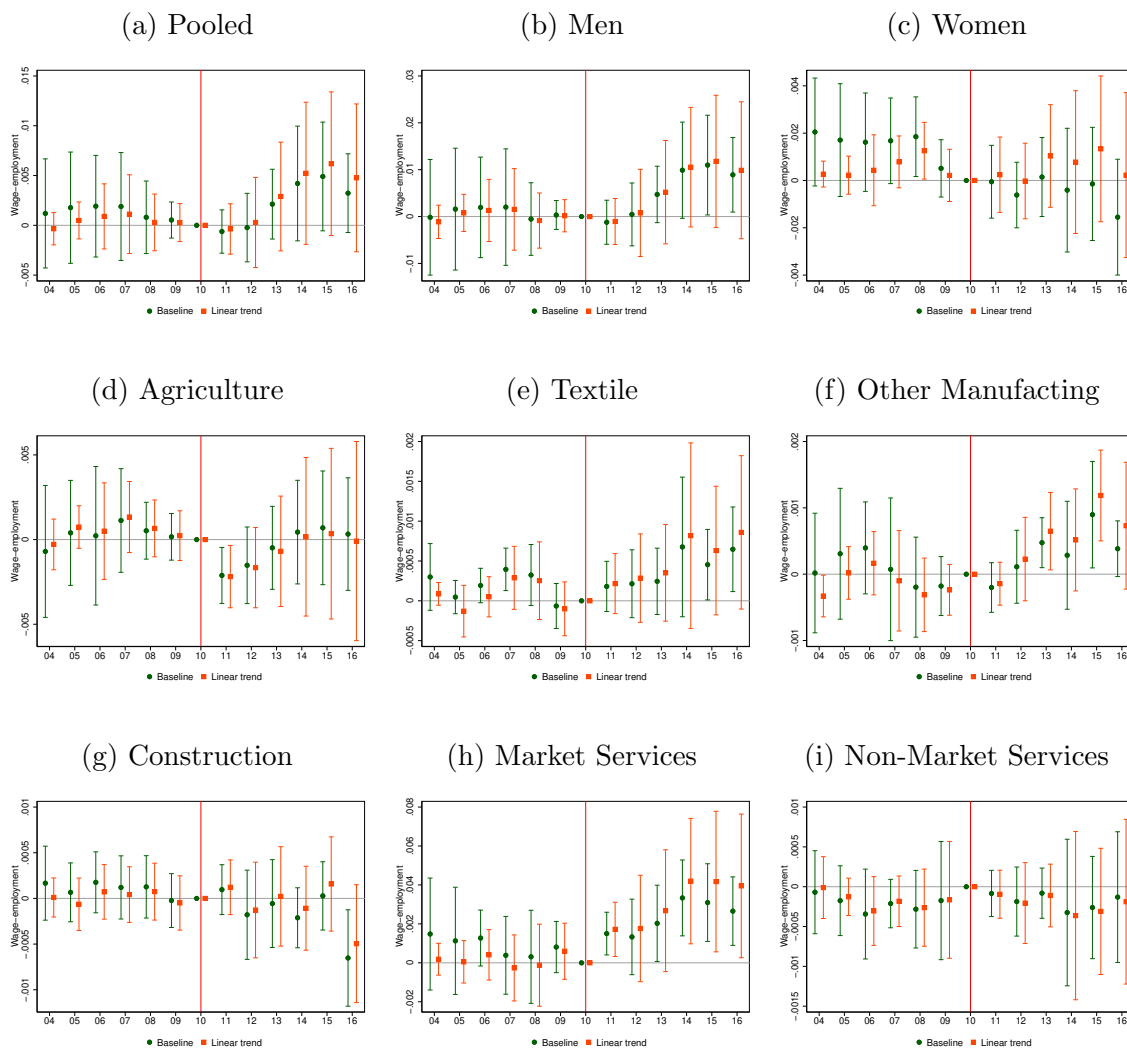
Figure B.3: Event-study figures for low-skill formal wage employment estimates in Figure 5



Notes: The reduced-form estimates come from the regression  $y_{p,t} = \sum_{j \neq 2009} \beta_j (\text{year}_j \times \tilde{Z}_p) + \delta_p + \delta_t^s + \epsilon_{p,t}$ , where  $\tilde{Z}_p$  is the standardized version of the instrument. In the baseline specification, the outcome variable is the raw wage employment rate. In the “linear trend” specification, I use the detrended outcome, where I estimate the trend in a regression using data between 2004-2010:  $y_{p,t} = f_p + f_t + y_{tr,p} * t + \xi_{p,t}$ . The industry specifications follow the ISIC standards. Standard errors are clustered at nuts2 region level. The 95% confidence interval is plotted.



Figure B.4: Event-study figures for low-skill formal non-wage employment estimates in Figure 7



Notes: The reduced-form estimates come from the regression  $y_{p,t} = \sum_{j \neq 2009} \beta_j (\text{year}_j \times \tilde{Z}_p) + \delta_p + \delta_t^s + \epsilon_{p,t}$ , where  $\tilde{Z}_p$  is the standardized version of the instrument. In the baseline specification, the outcome variable is the raw wage employment rate. In the “linear trend” specification, I use the detrended outcome, where I estimate the trend in a regression using data between 2004-2010:  $y_{p,t} = f_p + f_t + y_{tr,p} * t + \xi_{p,t}$ . The industry specifications follow the ISIC standards. Standard errors are clustered at nuts2 region level. The 95% confidence interval is plotted.

## C Alternative Identification Strategies

As described in the introduction, several papers investigated the effects of the Syrian refugees on the Turkish labor markets. Using different identification strategies, this literature mostly found confounding results, especially on formal employment. Del Carpio and Wagner (2015); Ceritoglu et al. (2017); Aksu et al. (2022) all document a decline in informal employment among natives as a consequence of the refugee shock, which is likely the only unchallenged result in this literature. Del Carpio and Wagner (2015) find an increase in formal employment, but only for low-skill men. However, using very similar data Akgündüz and Torun (2020) claim instead that high-skill employment (which is mostly formal) has increased. Across men and women, Aksu et al. (2022) argue that refugees lead to an increase in formal employment for men, and a decrease for women. Their results are challenged by Erten and Keskin (2021), who find a decrease in employment only for women and not for men. Using a generalized synthetic control method to adjust for pre-trends, Cengiz and Tekgüç (2022) claim that there was no employment loss among natives due to the refugee shock. On firm entry, Altındağ et al. (2020) show that the number of firms created by Syrians increased, but that the increase in total new firms was limited. In contrast, Akgunduz et al. (2022) find a sizeable increase in the number of exporter firms.

I argue in the paper that these opposing findings on native employment result from a combination of (1) not separating employment into components that are governed by different economic forces (such as wage and non-wage employment), and (2) not adjusting for non-linear trends in the IV-DiD design. Here, I provide support to my latter claim. I investigate the different identification strategies used in these papers and show in which ways they fail to account for the pre-trends in the data.

First, the first set of papers studying the Syrian refugee crisis in Turkey worked with limited datasets. Since the crisis began in 2011, they used data starting from 2009-2010 to make inference about the causal effect of refugees. However, as I have shown in figures B.1, this approach fails to account for the differential pre-trends between the regions close to the border and regions further away, especially for low-skill employment in the informal and formal sectors. Therefore, their estimates do not capture the causal effect of Syrian refugees.

Second, a common approach in the literature is to control for aggregate region-year fixed effects. Turkish Statistical Institute follows the NUTS classification (Nomenclature of territorial units for statistics) for data collection. Turkey has 12 nuts1 regions, 26 nuts2 regions, and 81 nuts3 regions (provinces). For instance, employment statistics are gathered at nuts2 level. A relatively common approach in this literature to deal with the pre-trend problem is to control for region-year fixed effects. In practice, researchers have used either

12\*(number of years) nuts1-year fixed effects, or they created another 5-region specification and controlled for 5-region-year fixed effects.

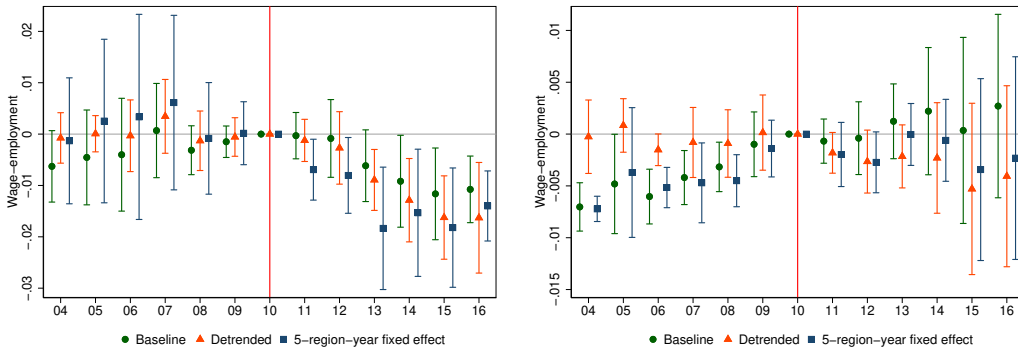
This approach has two weaknesses. First, the main idea behind the distance instrument is that the distance-induced variation in the treatment is an exogenous variation. Region-year fixed effects absorb a significant portion of this exogenous variation, and relies on the variation in treatment within a region for identification. This variation is arguably more likely to suffer from endogeneity. For instance, imagine two provinces in the same region that are equally distant to the border, but one province receives a larger treatment than the other. Generally, we would argue that comparing these two regions may not be a good idea to estimate the causal effect due to endogenous sorting of refugees. However, if we were to control for region-year fixed effect in a regression, we would rely solely on the treatment variation within these two provinces for identification, the outcome that we wanted to avoid in the first place. Put differently, the controlling for region-year fixed effects in a IV-DiD design in which distance is the instrument *changes* the identifying assumptions substantially. This is less of a statistical argument and more of an econometric one. Even if pre-trends looked good, we could still have doubts about the validity of the parallel trends assumption.

Besides the economic arguments, from a statistical perspective controlling for region-year fixed effects does not account for the pre-trends, and actually exacerbates the level of pre-trends for certain outcomes. To show this, I estimate the same event-study figures I showed in the main text, but this time controlling for 5-region-year time fixed effects, and 12-region-year time fixed effects. To allow for an easy comparison with the results from the main text, in each figure I show the event-study estimates using (1) a baseline model in which I do not make any adjustment for pre-trends, (2) detrending the outcome in the pre-period, the preferred strategy in the main text, and (3) controlling for region-year fixed effects. I plot the estimates on Figures C.5 and C.6.

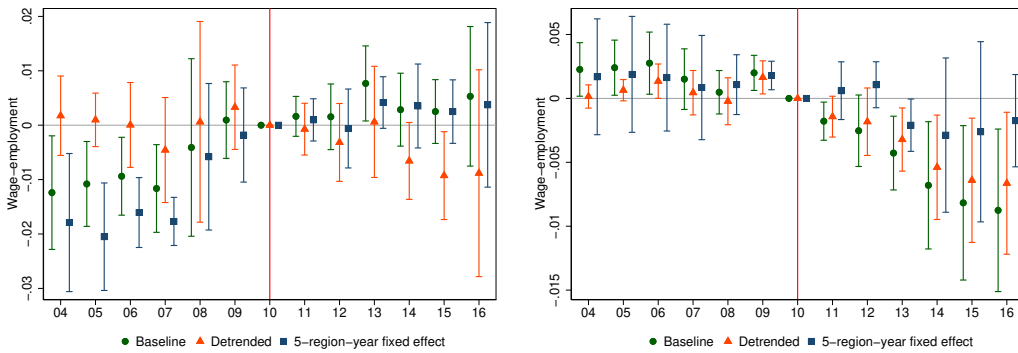
Another set of alternatives has been to control for linear trends at the 5-region, 12-region (nuts1), and 26-region (nuts2) level within the 2SLS regression. These methodologies are similar in spirit to my preferred methodology. The key differences are two-folds. First, 5-region and 12-region level linear trends do not seem to account for the non-linear pre-trends as much. This is visible in Figures C.8, C.9, and C.10 which show that the pre-trends mostly remain robust to these controls. Second, controlling for region-specific linear trends inside the 2SLS regression means that the same controls are included in the first-stage. Because the treatment takes a value of zero for half of the time, and gets a positive value in the other half of the time, a linear trend using the full timespan generates residual treatment that is nonzero in the pre-period. Consequently, the instrument starts predicting a pseudo-treatment in the pre-period even though there should be none. This can be seen in Figure

Figure C.5: Event-study figures with 5-region-year f.e. and detrended data

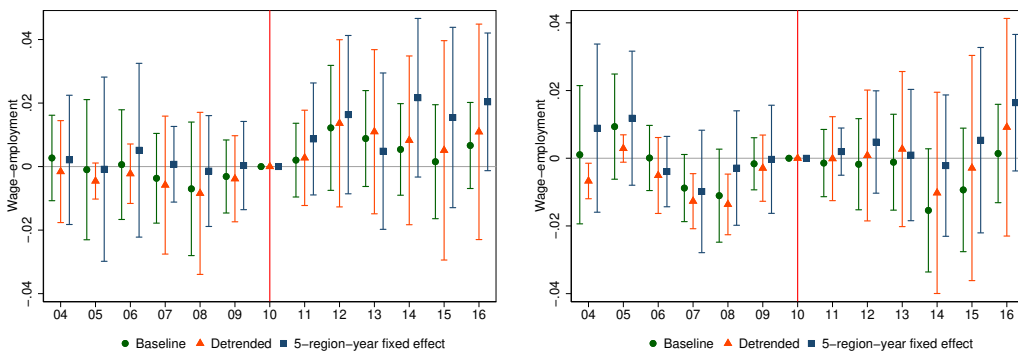
(a) Informal wage-employment



(b) Formal low-skill wage-employment



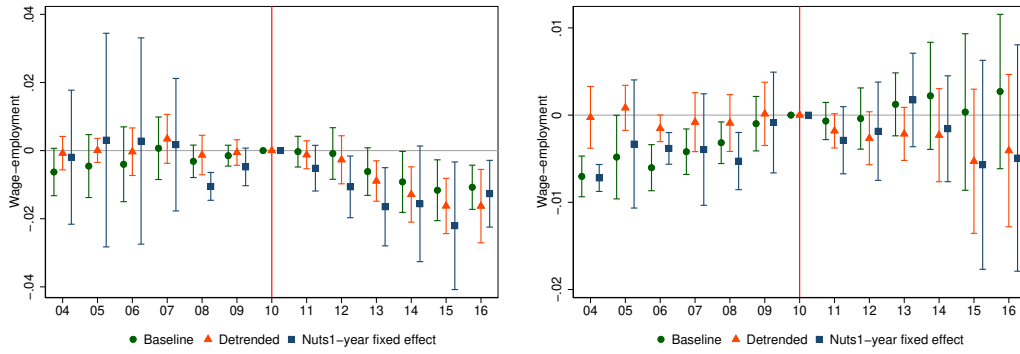
(c) Formal high-skill wage-employment



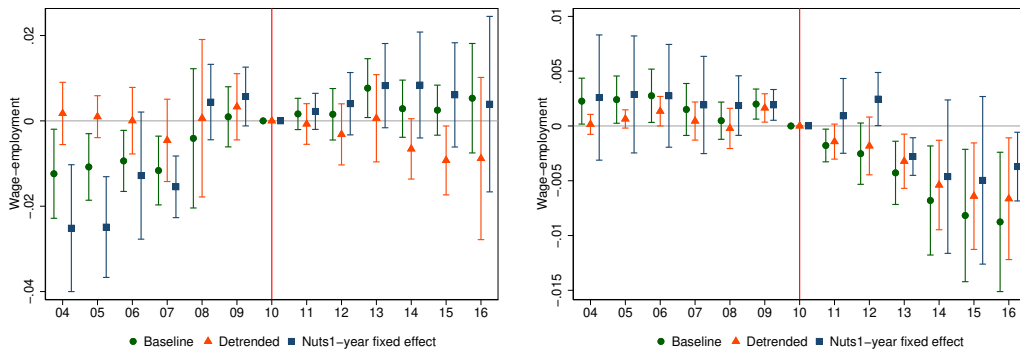
Notes: This figure shows the event-study estimates from three different identification strategies: the baseline DiD model, the preferred detrending method employed in the main text, and an alternative commonly used in the literature: controlling for 5-region-year fixed effects.

Figure C.6: Event-study figures with nuts1-year f.e. and detrended data

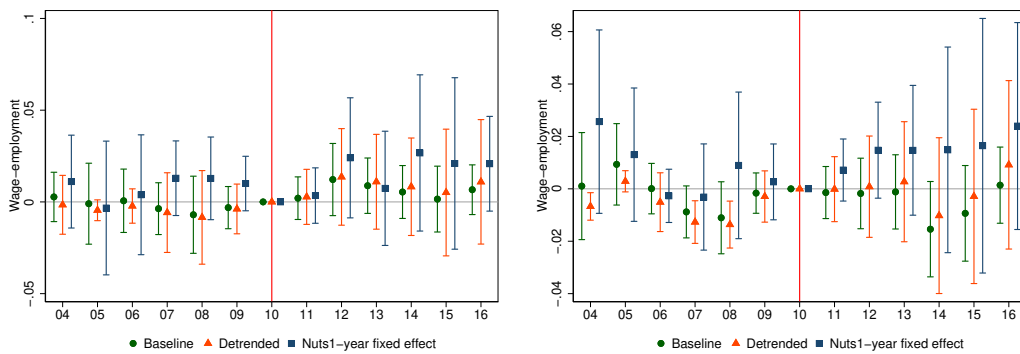
(a) Informal wage-employment



(b) Formal low-skill wage-employment

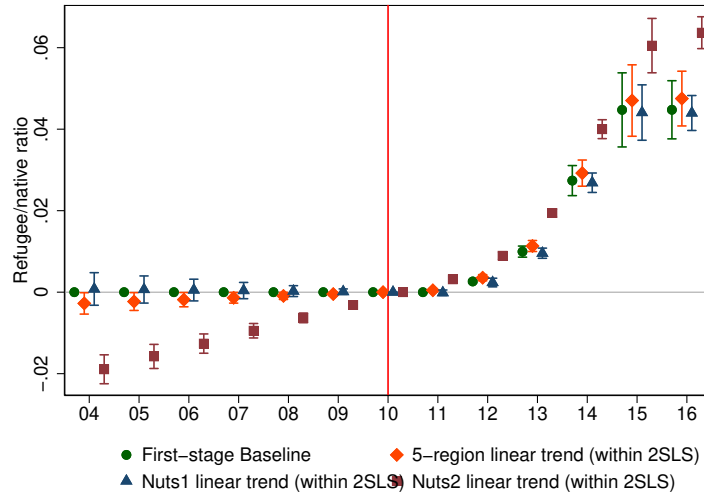


(c) Formal high-skill wage-employment



Notes: This figure shows the event-study estimates from three different identification strategies: the baseline DiD model, the preferred detrending method employed in the main text, and an alternative commonly used in the literature: controlling for nuts1-year fixed effects.

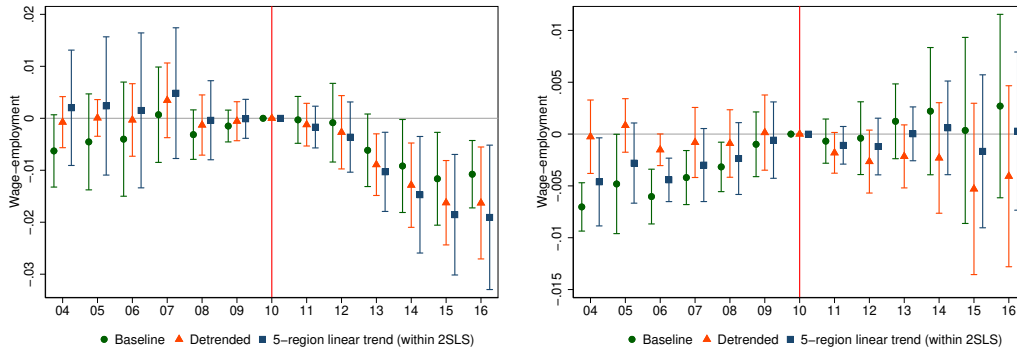
Figure C.7: First-stage under different linear-trend adjustments



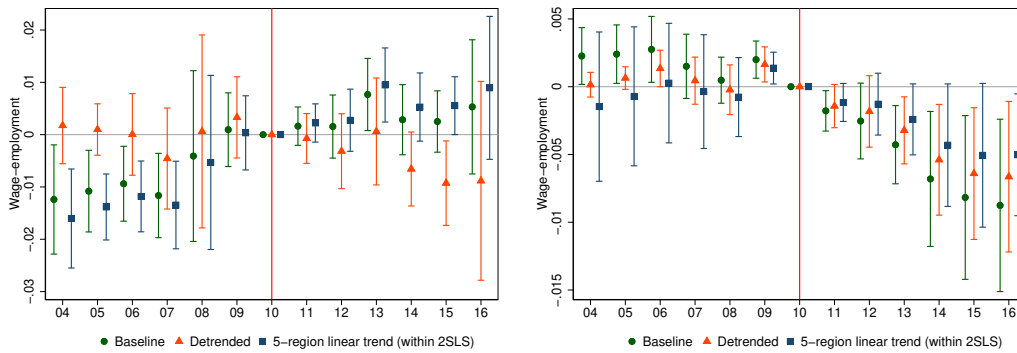
C.7, in which I show the first-stage estimates that the instrument predicts at each period while adjusting for different levels of linear trends. Ex-post, controlling for 5-region linear trends creates a tiny shadow pseudo-treatment in the pre-period. However, controlling for nuts2-linear trends (meaning I fit a separate linear trend for each unit in my study) creates substantial pseudo-treatment in the pre-period.

Figure C.8: Event-study figures with nuts1-year f.e. and detrended data

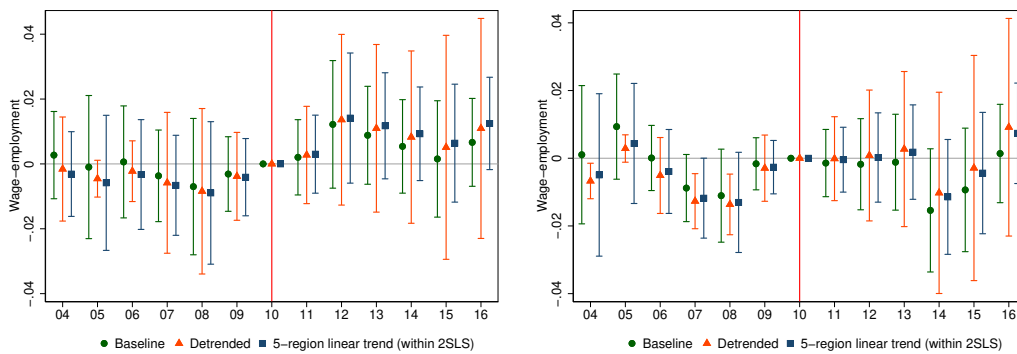
(a) Informal wage-employment



(b) Formal low-skill wage-employment



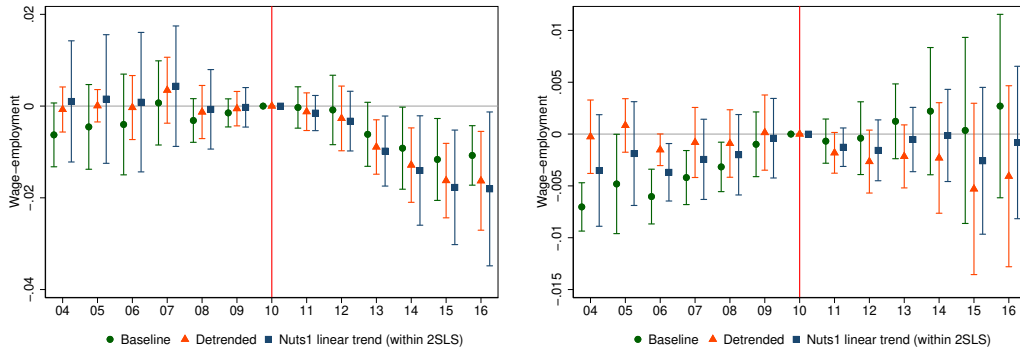
(c) Formal high-skill wage-employment



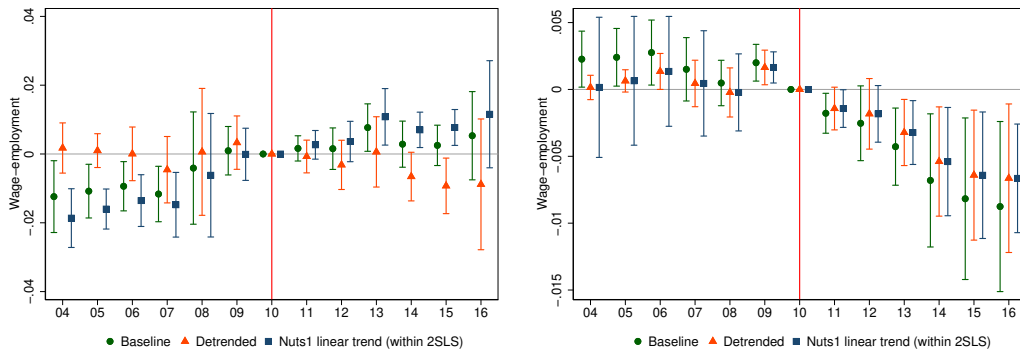
Notes: This figure shows the event-study estimates from three different identification strategies: the baseline DiD model, the preferred detrending method employed in the main text, and an alternative commonly used in the literature: controlling for linear trends at 5-region level.

Figure C.9: Event-study figures with nuts1-year f.e. and detrended data

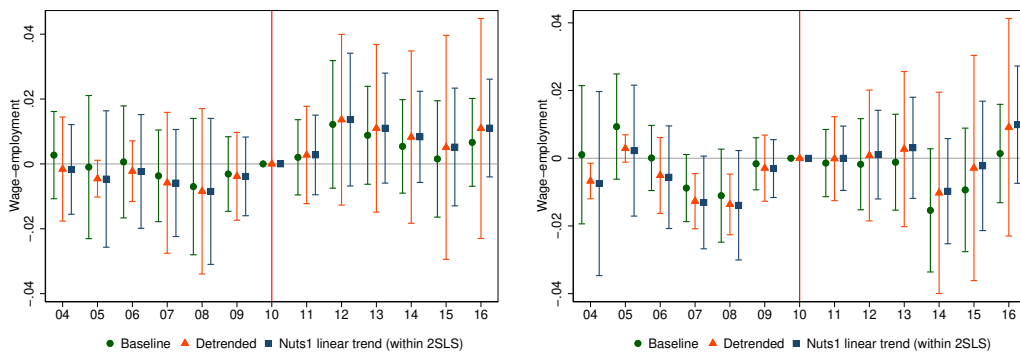
(a) Informal wage-employment



(b) Formal low-skill wage-employment



(c) Formal high-skill wage-employment

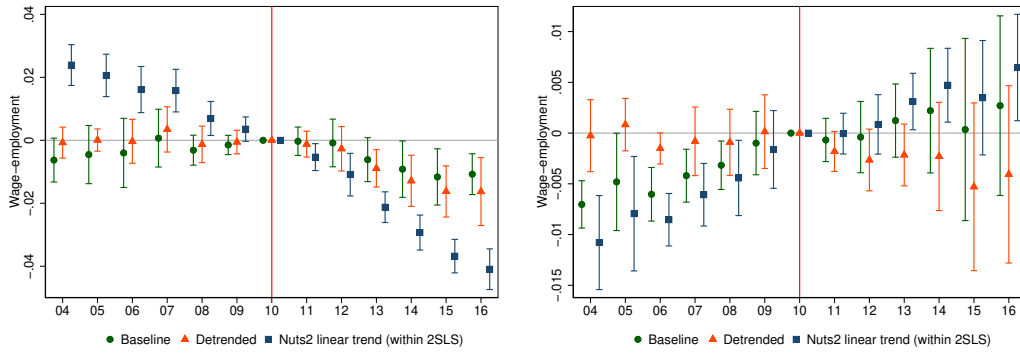


Notes: This figure shows the event-study estimates from three different identification strategies: the baseline DiD model, the preferred detrending method employed in the main text, and an alternative commonly used in the literature: controlling for linear trends at nuts1 level.

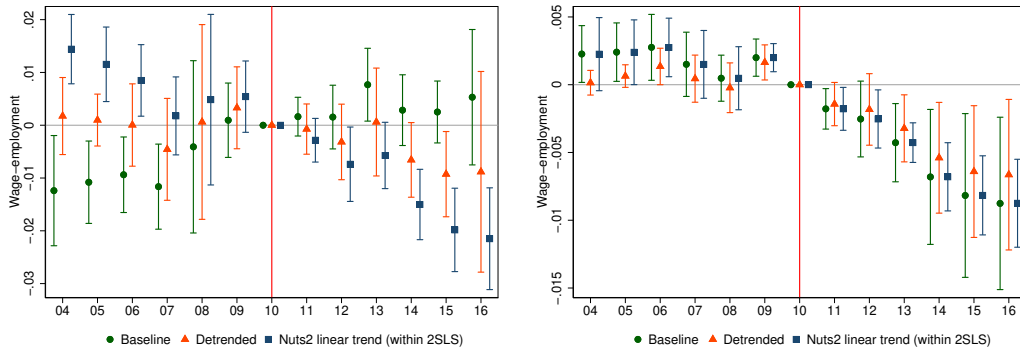


Figure C.10: Event-study figures with nuts2 linear trends and detrended data

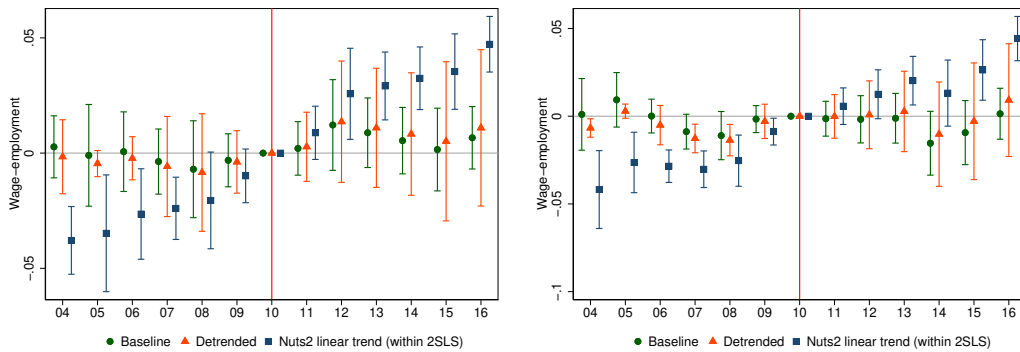
(a) Informal wage-employment



(b) Formal low-skill wage-employment



(c) Formal high-skill wage-employment



Notes: This figure shows the event-study estimates from three different identification strategies: the baseline DiD model, the preferred detrending method employed in the main text, and an alternative commonly used in the literature: controlling for linear trends at nuts2 level.

## D Refugees' effect on native population

In Turkey, population and number of new firms is strongly correlated. Figure [D.11a](#) plots the natural logarithm of the number of new firms and native population at the province level in 2009. There is strong correlation between new entrants and local population. A linear line fits the data almost perfectly with an R-square of 0.94. The elasticity (at the cross-section) is around 1.1. Of-course, this does not imply causation: cities where many people live may have other amenities that allow for new firm formation. Within province variation in population and firm entry is more informative. Regressing (log of) number of new firms on (log of) population while controlling for province and year fixed effects in the pre-period show that the true elasticity is around 0.75, which is still large.

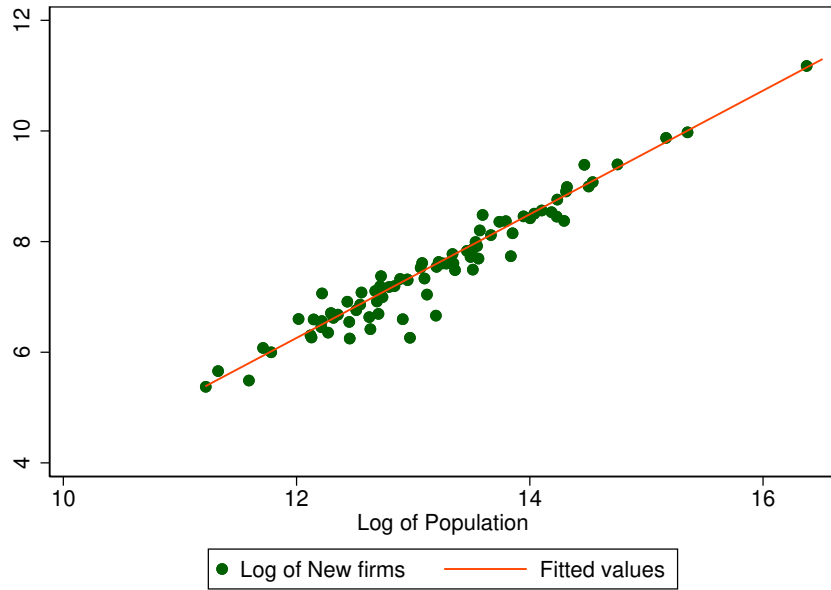
This relationship between population and firm entry continues in the post period for the control provinces that did not receive many refugees. Figure [D.11b](#) shows the timeseries evolution of firm entry and total population (refugees and natives) in Ankara, the capital of Turkey. By the end of 2016, the number of refugees in Ankara was only 1.2% of the native population, a relatively small treatment intensity. Over time, both the population and number of new firms continue increasing in a steady fashion. The correlation between population and firm entry remains positive.

In contrast, Figure [D.11c](#) plots the same statistics for Kilis, the most intensely treated province in Turkey. By 2015, the number of Syrian refugees in Kilis was 94% of the native population. The total population practically doubled between 2011 and 2016. Despite this large increase in population, the number of new firms did not increase. Except for a one time increase in new firms in 2012, the average number of new firms remain constant in the pre and post periods.

These figures are suggestive that the refugees do not lead to an increase in the number of new firms despite increasing the population, but more credible results require a regression analysis. For instance, refugees may affect the native population (by changing in and out-migration), which might lead to a decline in new firms. After all, Turkish natives are richer and more educated than the Syrian refugees, hence the elasticity of new formation with respect to native population should be different than that the same elasticity w.r.t. refugees. Hence, understanding the causal effect of refugees on native population is crucial for inference on firm entry.

Following the structure in Section 4, I estimate the reduced-form effect of the instrument on native population, in and out migration of natives at the province level. I plot the results on Figure [D.12](#). On figure [D.12a](#), we see that the provinces closer to the border observed statistically significant changes in both in-migration and out-migration. The effects are

Figure D.11: Relationship between number of new firms, population and refugees



(a) Market size and firm entry in 2009

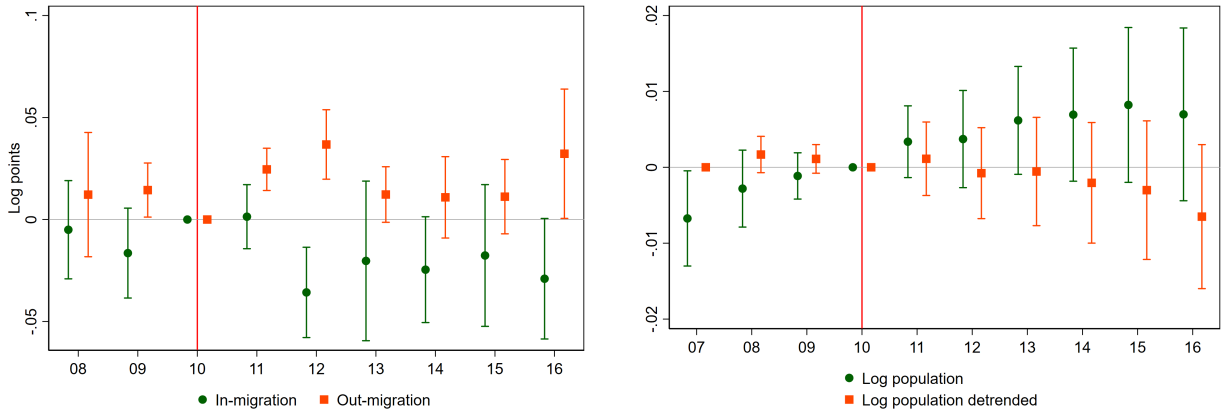


(b) Population and firm entry in Ankara (control)



(c) Population and firm entry in Kilis (treated)

Figure D.12: Event-study figures on native population



(a) RF effect on native in-out migration

(b) RF effect on native population

apparent initially in 2011 and 2012 when the Syrian war began (even before refugees started coming in masses), but then subside until the end of 2015, and then slightly increase again in 2016.

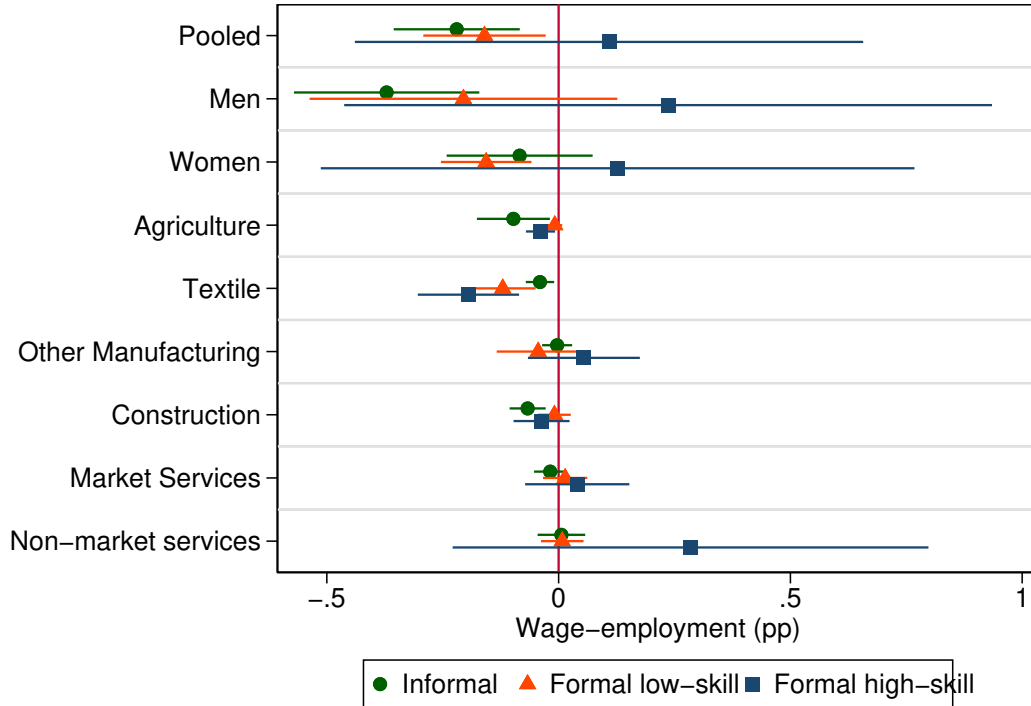
Overall, it is apparent that the instrument does capture some statistically significant changes in native in-migration and out-migration. However, these effects are small in magnitude. For instance, a 1 standard deviation in the predicted treatment intensity increases (decreases) out-migration (in-migration) by less than 3%. Whereas this may sound large, in/out-migration each constitutes around 3% of the native population in the more intensely treated provinces in each year. Hence, a 2 standard deviation increase in treatment intensity decreases native population in a province by around 0.36%. Given the 0.75 elasticity between firm entry and native population, this would lead only to a mild 0.27% decrease in the number of new firms.

In fact, the changes in in and out migration does not lead to a detectable change in native population. Figure D.12b plots the same event-study figure on native population. There is a clear pre-trend in the data. Regions closer to the border have higher birth-rates, and hence their native population grows faster than elsewhere. I adjust for this difference by estimating a linear trend in the pre-period and assuming it continues in the post. The figure does detect a mild decline in native population in regions close to the border, but this decline is statistically insignificant and small in magnitude, as I had predicted based on the mild effects on in and out migration. The 2SLS estimates, which I show in columns 1–3 of Table D.3, show that a 1pp increase in refugee native ratio decreases native population by only 0.06%, in-migration by only 0.17%, and increases out-migration by only 0.05%. All these estimates are statistically insignificant.

Table D.3: 2SLS estimates of refugees' effect on native population, in-out migration, and formal entry

	(1)	(2)	(3)	(4)	(5)
	ln(pop_native)	ln(in-migration)	ln(out-migration)	ln(entry)	ln(entry)
refugee/native	-0.0644 (0.0572)	-0.170 (0.184)	0.0538 (0.120)	0.0699 (0.106)	-0.925*** (0.277)
Controls					
region	Yes	Yes	Yes	Yes	Yes
year	Yes	Yes	Yes	Yes	Yes
log(population)	-	-	-	No	Yes
N	810	729	729	648	648

Figure E.13: Effect of Refugees on native wage-employment rates



Notes: The 2SLS estimates come from the regression  $\tilde{y}_{p,t} = \beta R_{p,t} + \delta_p + \delta_t + \epsilon_{p,t}$ , where I instrument for the refugee/native ratio  $R_{p,t}$  with the distance instrument  $Z_p$ .  $\tilde{y}_{p,t} = y_{p,t}^s - (t - 2004) * y_{tr,p}^s$  is the detrended outcome, where I estimate the trend in a regression using data between 2004-2010:  $y_{p,t} = f_p + f_t + y_{tr,p} * t + \xi_{p,t}$ . The outcome variables are (1) informal wage-employment rate, (2) formal wage-employment rate for the low-skilled, and (3) formal wage-employment rate for the high-skilled. The first row shows the estimates on the pooled data. The second and third rows condition on men and women separately. Fourth row considers the employment rate only in the high-intensity sectors, which are manufacturing, construction, and agriculture. This changes the numerator, not the denominator, in the calculation of the employment rate. The tradable industries are Manufacturing, Mining, and Agriculture. Standard errors are clustered at nuts2 region level. The 95% confidence interval is plotted.

## E Additional empirical Checks

Table E.4: Refugees' effect on the employment rate of natives

All industries		Agriculture	Textile	Other Man- ufacturing	Construction	Market Ser- vices	Non-market Services	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Informal	Pooled	-0.381 (0.302)	-0.310 (0.235)	-0.052*** (0.021)	0.012 (0.022)	-0.063*** (0.022)	0.000 (0.051)	0.033 (0.037)
	Men	-0.301 (0.265)	-0.203 (0.185)	-0.052*** (0.012)	0.030 (0.029)	-0.133*** (0.046)	0.035 (0.085)	0.022 (0.024)
	Women	-0.460 (0.364)	-0.406 (0.296)	-0.052 (0.032)	-0.006 (0.019)	0.000 (0.003)	-0.039 (0.025)	0.043 (0.064)
Low-skill Formal	Pooled	-0.037 (0.104)	-0.005 (0.055)	-0.104*** (0.040)	-0.021 (0.045)	-0.014 (0.017)	0.102* (0.052)	0.005 (0.024)
	Men	0.051 (0.196)	-0.021 (0.102)	-0.196** (0.090)	-0.005 (0.088)	-0.029 (0.038)	0.27*** (0.103)	0.032 (0.053)
	Women	-0.148*** (0.059)	0.004 (0.026)	-0.04*** (0.016)	-0.045** (0.020)	-0.002 (0.003)	-0.047** (0.020)	-0.018** (0.008)
High-skill Formal	Pooled	0.221 (0.346)	-0.054 (0.035)	-0.185*** (0.055)	0.094 (0.067)	-0.041 (0.029)	0.113 (0.082)	0.293 (0.276)
	Men	0.492 (0.478)	-0.063 (0.057)	-0.244*** (0.083)	0.185** (0.087)	-0.061 (0.052)	0.306*** (0.110)	0.369 (0.316)
	Women	0.108 (0.340)	-0.023* (0.012)	-0.032 (0.034)	0.039 (0.053)	0.018 (0.017)	-0.057 (0.074)	0.162 (0.303)

The 2SLS estimates come from the regression  $\tilde{y}_{p,t} = \beta R_{p,t} + \delta_p + \delta_t + \epsilon_{p,t}$ , where I instrument for the refugee/native ratio  $R_{p,t}$  with the distance instrument  $Z_{p,t} \cdot \tilde{y}_{p,t}^s = y_{p,t}^s - (t - 2004) * y_{tr,p}^s$  is the detrended outcome, where I estimate the trend in a regression using data between 2004-2010:  $y_{p,t} = f_p + f_t + y_{tr,p} * t + \xi_{p,t}$ . Standard errors are clustered at the region level. Industry codes are determined according to ISIC. Details can be found using this link: <https://ilosstat.ilo.org/resources/concepts-and-definitions/classification-economic-activities/>

Table E.5: Refugees' effect on the wage employment rate of natives

	All industries (1)	Agriculture (2)	Textile (3)	Other Manufacturing (4)	Construction (5)	Market Services (6)	Non-market Services (7)
Pooled	-0.22*** (0.069)	-0.098*** (0.040)	-0.04*** (0.016)	-0.003 (0.017)	-0.067*** (0.020)	-0.018 (0.018)	0.006 (0.026)
Informal							
Men	-0.371*** (0.102)	-0.117* (0.066)	-0.067*** (0.017)	0.001 (0.027)	-0.14*** (0.042)	-0.022 (0.038)	-0.027 (0.030)
Women	-0.084 (0.080)	-0.081*** (0.030)	-0.015 (0.020)	-0.007 (0.013)	0.000 (0.002)	-0.017 (0.014)	0.036 (0.048)
Pooled	-0.16*** (0.067)	-0.008 (0.008)	-0.121*** (0.036)	-0.044 (0.046)	-0.009 (0.018)	0.014 (0.024)	0.008 (0.023)
Low-skill							
Formal							
Men	-0.205 (0.169)	-0.009 (0.017)	-0.23*** (0.079)	-0.059 (0.090)	-0.020 (0.040)	0.076 (0.047)	0.036 (0.050)
Women	-0.156*** (0.050)	-0.009** (0.004)	-0.041*** (0.015)	-0.043** (0.020)	-0.001 (0.002)	-0.047*** (0.019)	-0.016 (0.011)
Pooled	0.109 (0.280)	-0.039*** (0.016)	-0.194*** (0.056)	0.055 (0.062)	-0.037 (0.031)	0.040 (0.057)	0.285 (0.262)
High-skill							
Formal							
Men	0.236 (0.356)	-0.054** (0.025)	-0.256*** (0.075)	0.114 (0.077)	-0.061 (0.051)	0.15** (0.069)	0.343 (0.295)
Women	0.127 (0.327)	-0.014** (0.007)	-0.038 (0.034)	0.038 (0.054)	0.019 (0.017)	-0.056 (0.065)	0.178 (0.295)

The 2SLS estimates come from the regression  $\tilde{y}_{p,t} = \beta R_{p,t} + \delta_p + \delta_t + \epsilon_{p,t}$ , where I instrument for the refugee/native ratio  $R_{p,t}$  with the distance instrument  $Z_{p,t}$ .  $\tilde{y}_{p,t} = y_{p,t}^s - (t - 2004) * y_{tr,p}^s$  is the detrended outcome, where I estimate the trend in a regression using data between 2004-2010:  $y_{p,t} = f_p + f_t + y_{tr,p} * t + \xi_{p,t}$ . Standard errors are clustered at the region level. Industry codes are determined according to ISIC. Details can be found using this link: <https://hlostat.ilo.org/resources/concepts-and-definitions/classification-economic-activities/>

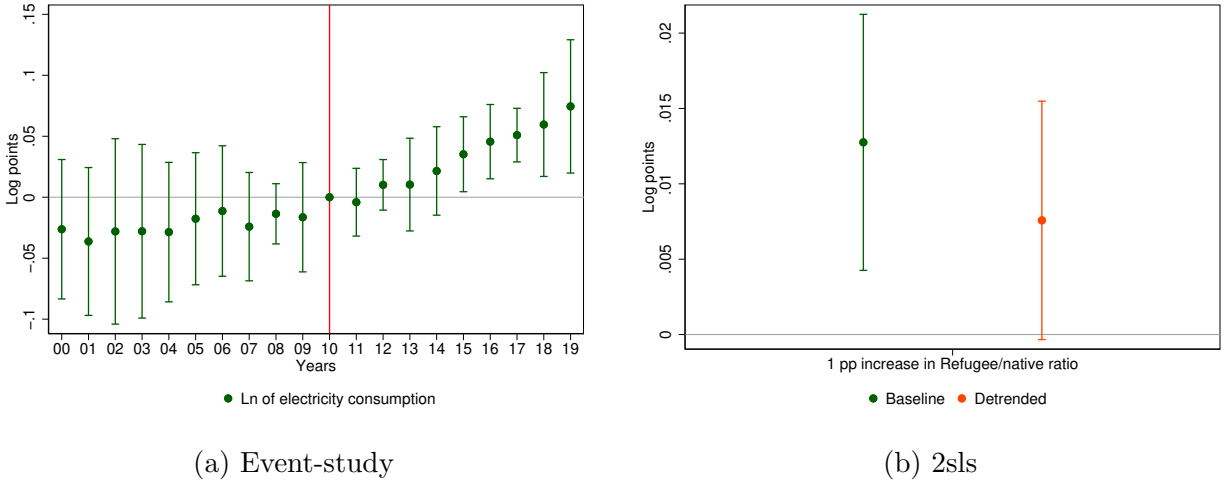


Table E.6: Refugees' effect on the non-wage employment rate of natives

All industries		Agriculture	Textile	Other Man- ufacturing	Construction	Market Ser- vices	Non-market Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Informal	Pooled	-0.381 (0.302)	-0.052*** (0.021)	0.012 (0.022)	-0.063*** (0.022)	0.000 (0.051)	0.033 (0.037)
	Men	-0.301 (0.265)	-0.052*** (0.012)	0.030 (0.029)	-0.133*** (0.046)	0.035 (0.085)	0.022 (0.024)
	Women	-0.460 (0.364)	-0.052 (0.032)	-0.006 (0.019)	0.000 (0.003)	-0.039 (0.025)	0.043 (0.064)
Low-skill Formal	Pooled	-0.037 (0.104)	-0.104*** (0.040)	-0.021 (0.045)	-0.014 (0.017)	0.102* (0.052)	0.005 (0.024)
	Men	0.051 (0.196)	-0.196** (0.090)	-0.005 (0.088)	-0.029 (0.038)	0.27*** (0.103)	0.032 (0.053)
	Women	-0.148*** (0.059)	-0.04*** (0.016)	-0.045** (0.020)	-0.002 (0.003)	-0.047** (0.020)	-0.018** (0.008)
High-skill Formal	Pooled	0.221 (0.346)	-0.185*** (0.055)	0.094 (0.067)	-0.041 (0.029)	0.113 (0.082)	0.293 (0.276)
	Men	0.492 (0.478)	-0.244*** (0.083)	0.185** (0.087)	-0.061 (0.052)	0.306*** (0.110)	0.369 (0.316)
	Women	0.108 (0.340)	-0.023* (0.012)	0.039 (0.053)	0.018 (0.017)	-0.057 (0.074)	0.162 (0.303)

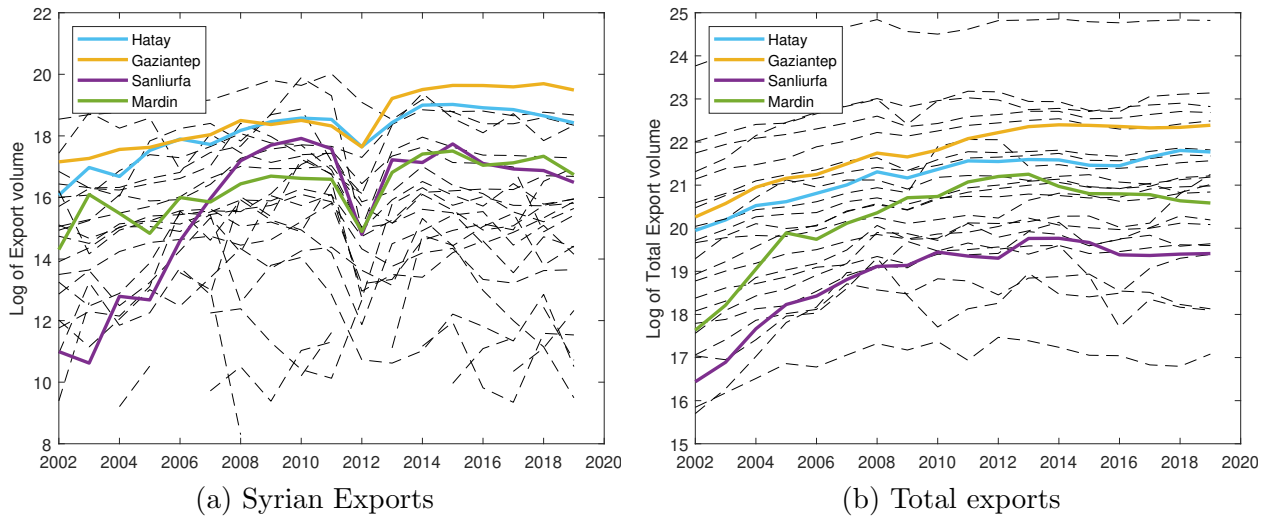
The 2SLS estimates come from the regression  $\tilde{y}_{p,t} = \beta R_{p,t} + \delta_p + \delta_t + \epsilon_{p,t}$ , where I instrument for the refugee/native ratio  $R_{p,t}$  with the distance instrument  $Z_{p,t} \cdot \tilde{y}_{p,t}^s = y_{p,t}^s - (t - 2004) * y_{tr,p}^s$  is the detrended outcome, where I estimate the trend in a regression using data between 2004-2010:  $y_{p,t} = f_p + f_t + y_{tr,p} * t + \xi_{p,t}$ . Standard errors are clustered at the region level. Industry codes are determined according to ISIC. Details can be found using this link: <https://ilo-stat.lib.org/resources/concepts-and-definitions/classification-economic-activities/>

Figure E.14: Effect of refugees on electricity consumption



Notes: Both figures uses electricity consumption data of Turkstat. For consistency with the rest of the paper, I collapse the data at the 26 region level. The 2SLS estimates use data until 2016 for consistency with the rest of the paper. 95% confidence intervals are shown. Standard errors are clustered at the NUTS-2 region level

Figure E.15: Timeseries of Export volume across regions



Notes: Hatay, Gaziantep, Sanliurfa and Mardin regions are the four regions (at NUTS2 level) at the Syrian border.

## F Derivations of the baseline model

To calculate these elasticities, first take the logarithm of the FOCs:

$$\begin{aligned}(\rho - 1 - \gamma)\log L_i &= \log w_i + \log(1 + \gamma) - \log(\alpha\eta) - \frac{\alpha - \rho}{\rho}\log(\eta L_i^\rho + (1 - \eta)L_f^\rho) \\(\rho - 1)\log L_f &= \log w_f + \log(1 + \tau_w) - \log(\alpha(1 - \eta)) - \frac{\alpha - \rho}{\rho}\log(\eta L_i^\rho + (1 - \eta)L_f^\rho)\end{aligned}\tag{12}$$

Fix  $w_f = \bar{w}_f$ , and differentiate w.r.t.  $w_i$

$$\begin{aligned}(\rho - 1 - \gamma)\epsilon_{L_i, w_i} &= 1 - (\alpha - \rho)[s_i\epsilon_{L_i, w_i} + s_f\epsilon_{L_f, w_i}] \\(\rho - 1)\epsilon_{L_f, w_i} &= -(\alpha - \rho)[s_i\epsilon_{L_i, w_i} + s_f\epsilon_{L_f, w_i}]\end{aligned}\tag{13}$$

where  $s_i = \frac{\eta L_i^\rho}{\eta L_i^\rho + (1 - \eta)L_f^\rho}$  and  $s_f = \frac{(1 - \eta)L_f^\rho}{\eta L_i^\rho + (1 - \eta)L_f^\rho}$  are the informal and formal share in the production. Two linearly independent equations with two unknowns can easily be solved analytically, which reveals:

$$\epsilon_{L_f, w_i} = \frac{(\alpha - \rho)s_i}{1 - \rho - (\alpha - \rho)s_f}\epsilon_{L_i, w_i}\tag{14}$$

and

$$\begin{aligned}\epsilon_{L_i, w_i} &= -\frac{1 - \rho - (\alpha - \rho)s_f}{(1 - \rho + \gamma)(1 - \rho) - (\alpha - \rho)[(1 - \rho + \gamma)s_f + (1 - \rho)s_i]} \\ \epsilon_{L_f, w_i} &= -\frac{(\alpha - \rho)s_i}{(1 - \rho + \gamma)(1 - \rho) - (\alpha - \rho)[(1 - \rho + \gamma)s_f + (1 - \rho)s_i]}\end{aligned}\tag{15}$$

## G Extended model to explain the results on formal firm entry

I have two sets of main empirical findings in the main text. First, on the native employment side, the low-skill informal refugee shock reduces native wage-employment not only in the informal sector, but also in the formal sector (for the low-skilled). Second, on the firm entry side, I show that the Syrian refugees increase the number of new exporter and importer firms, do not change the number of nontrader firms, and weakly decrease the number of small firms (sole proprietorships). In the text, I explained that this change in the productivity distribution of new formal firms can be indicative of less productive entrepreneurs to remain unregistered. In this section, I formalize the economic forces behind this claim in an equilibrium model where firms can exploit both the intensive and extensive margins of informality. The model is based on Ulyssea (2018)’s framework to capture intensive and extensive margins of informality, but also uses some intuition from Melitz (2003) to divide formal firms into trader and nontrader types.

I begin with a close economy to set notation and intuition. Firms are heterogenous and indexed by their individual productivity,  $\theta$ . They produce a homogenous good using labor as their only input.<sup>36</sup> Product and labor markets are competitive, and formal and informal firms face the same prices. For simplicity, I assume that workers are perfect substitutes. The main insights of the model carry over to a model with multiple skill types. Formal and informal employees perform the same tasks within the firm, and therefore are perfect substitutes. On the labor supply side, workers are endowed with either formal or informal labor. Hence, there are natives who can provide only informal labor, and there are natives who can provide only formal labor.<sup>37</sup>

### G.1 Firms

Both formal and informal firms have access to the same technology. Output of a given firm with productivity  $\theta$  is given by  $y(\theta, \ell) = \theta q(\ell)$ , where the function  $q(\cdot)$  is assumed to be increasing, concave, and twice continuously differentiable.

Informal firms are able to avoid taxes and labor costs, but face a probability of detection by government officials. This expected cost takes the form of an ad-valorem labor distortion

---

<sup>36</sup>By assuming a homogenous good, I abstract away from the demand effects of the refugee shock. I talk about potential extensions at the end of this section.

<sup>37</sup>This is a reduced-form simplification. One can allow natives to search for both formal and informal jobs, but with heterogeneous productivity in searching for formal jobs. Since I do not have data on transitions from unemployment to formal/informal employment, I won’t dive into the details of such a search model. However, the main insights from my model would carry over.

denoted by  $\tau_i(\ell)$ , which is assumed to be increasing and convex in firm's size ( $\tau_i', \tau_i'' > 0$ ).<sup>38</sup> Informal firms' profit function is given by:

$$\pi_i(\theta, w_i) = \max_{\ell} \{\theta q(\ell) - w_i \ell \tau_i(\ell)\} \quad (16)$$

where the price of the final good is normalized to one.

Formal incumbents must comply with taxes and regulations, but they can hire informal workers to avoid the costs implied by the labor legislation. For formal firms, informal and formal workers are perfect substitutes. The hiring costs of formal and informal workers differ due to (1) different wages (e.g., there can be a binding minimum wage for formal workers), and (2) institutional reasons: formal firms have to pay a constant payroll tax on formal workers, while they face an increasing and convex expected cost to hire informal workers, which is summarized by the convex function  $\tau_{fi}(\cdot)$ . The cost of hiring  $\ell$  informal workers is given by  $\tau_{fi}(\ell) + w_i \ell$ , while the cost of hiring  $\ell$  formal workers is  $(1 + \tau_w)w_f \ell$ , where  $\tau_w$  is the payroll tax.

Formal firms' profit function can be written as follows:

$$\pi_f(\theta, w_i, w_f) = \max_{\ell_i, \ell_f} (1 - \tau_y) [\theta q(\ell_i + \ell_f) - \tau_{fi}(\ell_i) - \ell_i w_i + (1 + \tau_w)w_f \ell_f] \quad (17)$$

where  $\tau_y$  denotes the corporate tax. Formal firms maximizing profits reveals the demand for formal labor as a function of informal wages  $w_i$ , formal wages  $w_f$ , and productivity  $\theta$ . The demand for informal workers come both from informal firms and formal firms.

Becoming a formal firm introduces the technology to hire workers formally with constant marginal costs as opposed to informally with increasing marginal costs. Hence, more productive firms that want to hire more workers become formal.

## G.2 Entry

There are two periods. In period 1, a large mass  $\mathcal{M}$  of potential entrants observe their productivity, which is distributed according to the cdf  $G$ . To enter either sector, firms must pay a fixed cost that is assumed to be higher in the formal sector:  $E_f > E_i$ . If firms enter either sector, they can hire labor to produce and sell the final good in period 2.

As there is only one period after entry, firm's value function assumes a clean form:

$$V_s(\theta, w_i, w_f) = \pi_s(\theta, w_i, w_f)$$

---

<sup>38</sup>These assumptions can be rationalized, for instance, by the fact that larger firms have a greater probability of being caught.

Figure G.16: ZPC and free-entry



Potential entrants choose between three options. They can choose not to enter and receive zero payoff, enter the informal sector by paying entry cost  $E_i$ , or enter the formal sector by paying  $E_f$ . Given the value functions, a potential entrant with productivity  $\theta$  decides to:

- enter into the formal sector if  $V_f(\theta, w_i, w_f) - E_f > \max\{V_i(\theta, w_i) - E_i, 0\}$
- enter into the informal sector if  $V_i(\theta, w_i) - E_i > \max\{V_f(\theta, w_i, w_f) - E_f, 0\}$
- not enter into either sector otherwise

If entry in both sectors is positive, the following entry-conditions must hold:

$$\begin{aligned} V_i(\bar{\theta}_i, w_i, w_f) &= E_i \\ V_f(\bar{\theta}_f, w_i, w_f) &= V_i(\bar{\theta}_f, w_i) + (E_f - E_i) \end{aligned} \tag{18}$$

where  $\bar{\theta}_i$  and  $\bar{\theta}_f$  are the productivity of firms that are at the margin of entering into informal and formal sectors, respectively. The least productive entrepreneurs with productivity  $\theta < \bar{\theta}_i$  choose not to enter. Firms with productivity  $\theta \in \bar{\theta}_i, \bar{\theta}_f$  are productive enough to make positive profits and prefer the informal sector over formal sector. The more productive firms with productivity  $\theta > \bar{\theta}_f$  want to hire many workers, which is too costly to do in the informal sector due to the convex costs of hiring. In this model, the ability to hire workers with constant marginal cost is the only reason why firms wish to become formal. The sorting of firms into no entry, informal entry, and formal entry brackets based on their productivity draws is plotted in Figure G.16. The mass of new formal firms is given by  $(1 - \bar{\theta}_f)\mathcal{M}$ .

### G.3 Equilibrium

To close the model, I need to specify the labor supply. Let  $L_i^{N,S}(w_i)$  and  $L_f^{N,S}(w_f)$  be the informal and formal labor supply curves of natives.<sup>39</sup> Since formal and informal workers are substitutes, the labor demand for workers in one sector depends on the wages in both sectors. In equilibrium, labor markets must clear: informal and formal wages are such that

<sup>39</sup>Labor supply curves being independent of the wages in the other sector comes from natives having either formal or informal labor endowment. Relaxing this assumption would not change the predictions of the model.

labor supply equals labor demand.

$$\begin{aligned} L_i^S(w_i) &= L_i^D(w_i, w_f) \\ L_f^S(w_f) &= L_f^D(w_i, w_f) \end{aligned} \tag{19}$$

To summarize, the equilibrium conditions are given by the following conditions: (i) in period 1, the zero profit cutoff and free entry conditions hold in both sectors; and (ii) in period 2, labor markets clear. Product market clearing comes freely due to the Walras' Law.

## G.4 Effects of an informal labor supply shock

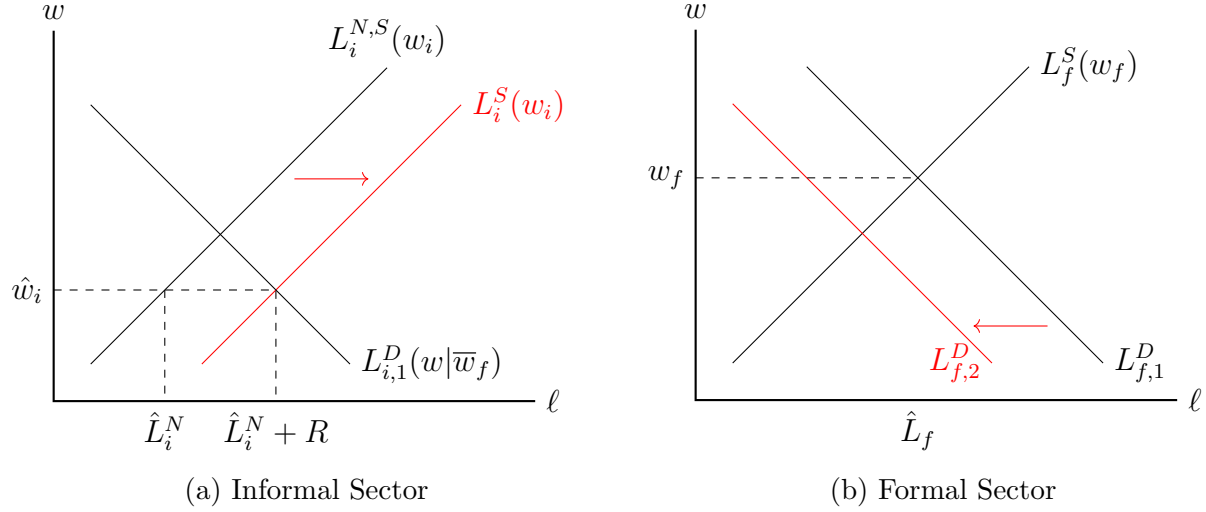
As in most refugee crises in the developing world, the overwhelming majority of Syrian refugees in Turkey did not have work permits. In the model, this will be captured by an increase in the informal labor supply. Figure G.17 shows how the refugee labor supply impacts the labor market equilibrium in this model. The left panel shows the equilibrium for informal workers and the right panel shows the equilibrium for formal workers. For ease of exposition, I assumed that refugees supply labor inelastically. This results in a parallel shift in the informal labor supply curve, causing (1) a decline in informal wages, (2) a decline in native informal employment, and (3) an increase in the aggregate informal employment. Since formal and informal workers are substitutes, the decrease in the informal wages incentivizes formal firms to rely more intensively on informal workers. This shifts the formal labor demand curve inward. As firms reduce their demand for formal workers, the amount of native formal employment decreases, despite refugees being unable to work formally.

On the extensive margin, I showed that despite a large increase in the population, which should have increased the number of new firms, refugees do not cause an increase in the number of new formal firms. I argue that this is due to informal refugees incentivizing the marginal new firms to remain informal instead. In this model, if access to informal workers is not easier for formal firms, it is easy to prove the following result:

**Proposition.** *The informal labor supply increase incentivizes firms to enter the informal sector instead of the formal sector. Formally,  $\frac{d\bar{g}_f}{dR} > 0$ , where  $R$  denotes the number of refugees in the economy.*

The intuition behind the proof is that the informal firm is more informal labor intensive. Hence, a decrease in wages for informal labor disproportionately increases the informal firm profits. Consequently, the marginal firm strictly prefers the informal sector as it provides

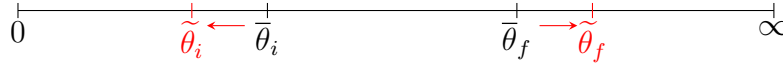
Figure G.17: Equilibrium with informal labor supply shock



Notes: For illustrative purposes, I assume the formal wage is fixed by a binding minimum wage. Otherwise, a decrease in the effective formal wage would also push the informal labor demand curve slightly upwards.

easier access to informal labor. This effect is visualised in figure G.18.<sup>40</sup>

Figure G.18: Effect of Informal LS on the extensive margin



This prediction of the model has significant implications regarding refugee crises. The current debate about the work permit status of refugees trades off the benefits of refugees becoming self-reliant (instead of relying on government resources) with native disemployment if refugees could work freely. This debate completely ignores the existence of an informal sector that absorbs the informal refugee labor supply. Taking firms' decision to be informal both on intensive and extensive margins rigorously reveals that by not allowing refugees to work formally, host countries are incentivizing firms to become more informal. This may have several implications, including decrease in tax revenue.

## G.5 Extensions

This framework can be extended to incorporate different skill groups in straight-forward fashion. For instance, one empirical finding of this paper is that the refugee shock has lead

<sup>40</sup>An untestable prediction of the model due to lack of data is that the decrease in informal wages should also increase the number of informal firms by allowing unproductive entrepreneurs to enter the informal sector instead of not creating any firm.



to an increase in wages of the high-skill workers in the formal sector. Model can formalize this finding by allowing formal firms to hire two types of labor, high and low skilled, which can enter the production function inside a CES aggregator. For example, let  $L_i, L_f, H$  denote the informal low-skill, formal low-skill and high-skill labor, respectively. The production function of a formal firm with productivity  $\theta$  can be written as:  $f(L, H) = \theta(\eta L^{\frac{\sigma-1}{\sigma}} + (1-\eta)H^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}$ , where  $L = L_i + L_f$  is the composite low-skill labor,  $\eta$  is the share parameter and  $\sigma$  is the elasticity of substitution between the two factors. As long as low-skill and high-skill labor are not perfect substitutes, for a given amount of H high-skill workers, hiring more low-skill labor makes high-skill labor more productive. If workers are paid their marginal product, this would increase the wages of the high-skilled workers.

Labor supply alone does not explain the entirety of my results. For example, I do not find a decrease in the number of new firms. On the contrary, I find a significant increase in the number of new firms that participate in international trade, especially those in the export market. The model and this empirical finding may seem inconsistent. On the contrary, the increase in new exporting firms helps discipline the type of firms the informalizing effects matter to. The key intuition is that the decrease in informal wages incentivizes the *marginal* firms to become informal. In contrast, a common finding in the trade literature is that more productive firms enter the exports market. To explain this finding, I marry Ulyssea (2018) and Melitz (2003) by giving formal firms the option to enter the exports market.

### G.5.1 Firms

Informal firms cannot participate in the exports market, and hence have to sell domestically for price  $p$ . Informal firms' profit function is now given by  $\pi_i(\theta, \ell_i) = \{p\theta q(\ell_i) - w_i \ell_i \tau_i(\ell_i)\}$ . Formal firms can participate in the export market. I assume a small, open economy where the local production or demand does not affect the international price  $\bar{p} > p$ , which is normalized to one. This simplifying assumption implies that for exporter firms, selling abroad is always more profitable than selling domestically. Consequently, non exporter firms sell only to domestic consumers, and exporter firms sell solely to international markets.<sup>41</sup> Hence, formal firms' profit function is given by:

$$\pi_f(\theta, \ell_i, \ell_f) = \begin{cases} p\theta q(\ell_i + \ell_f) - C(\ell_i, \ell_f) & \text{if non-exporter} \\ \theta q(\ell_i + \ell_f) - C(\ell_i, \ell_f) & \text{if exporter} \end{cases} \quad (20)$$

---

<sup>41</sup>Obviously, this is empirically false. I can get around this caveat by introducing a continuum of unique goods where producers value variety a la Melitz (2003), but this would introduce additional parameters to the model without adding much to the intuition.

where  $C(\ell_i, \ell_f) = \tau_{fi}(\ell_i) - \ell_i w_i + (1 + \tau_w)w_f \ell_f$  is the costs of hiring informal and formal workers. For notational simplicity, I will denote the profit function of the non-exporter and formal firm by  $\pi_f$  and that of the exporter firm by  $\pi_x(\theta)$ .

Introducing exporter firms serve two purposes. Mechanically, it introduces a second price that is set by the international markets, and hence unaffected by refugees. This enables me to model refugees' demand effect in a straight-forward way (Borjas, 2014). Second, it divides the set of (formal) entrepreneurs into two groups: those who are productive enough to export and others. This distinction helps separate the labor supply and entrepreneurial effects of refugees in a testable way, which will become apparent once I close the model.

### G.5.2 Entry

Entry is similar to the baseline model. There is a large mass  $\mathcal{M}$  of potential entrants who observe their productivity  $\theta \sim G$ . Entering the formal sector costs more than entering the informal sector:  $E_f > E_i$ . Additionally, becoming an exporter requires a fixed cost of entry a la Melitz (2003). Let  $E_x$  denote the total cost of becoming an exporter firm. Naturally,  $E_x > E_f$ .

As there is only one period after entry, firm's value function assumes a clean form  $V_s(\theta) = \pi_s(\theta, w_i, w_f)$ , where I suppress the wages in the value function for notational simplicity, and  $s \in \{i, f, x\}$ . Potential entrants choose between four options. They can choose not to enter and receive zero payoff, enter the informal sector by paying entry cost  $E_i$ , enter the formal sector as a non-exporter by paying  $E_f$ , or enter the exports market by paying  $E_x$ . Given the value functions, a potential entrant with productivity  $\theta$  decides to:

- enter into the export market if  $V_x(\theta) - E_x > \max\{V_f(p, \theta) - E_f, V_i(p, \theta) - E_i, 0\}$
- enter into the formal sector if  $V_f(p, \theta, w_i, w_f) - E_f > \max\{V_x(\theta) - E_x, V_i(p, \theta, w_i) - E_i, 0\}$
- enter into the informal sector if  $V_i(\theta, w_i) - E_i > \max\{V_x(\theta) - E_x, V_f(\theta, w_i, w_f) - E_f, 0\}$
- not enter into either sector otherwise

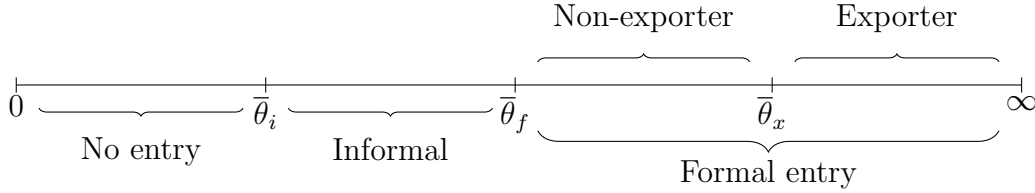
If entry in all sectors is positive, the following entry-conditions must hold:

$$\begin{aligned}
 V_i(\bar{\theta}_i) &= E_i \\
 V_f(\bar{\theta}_f) &= V_i(\bar{\theta}_f) + (E_f - E_i) \\
 V_x(\bar{\theta}_x) &= V_f(\bar{\theta}_x) - (E_x - E_f)
 \end{aligned} \tag{21}$$

where  $\bar{\theta}_i$ ,  $\bar{\theta}_f$ , and  $\bar{\theta}_x$  are the productivity of firms that are at the margin of entering into informal, formal, and exporter sectors, respectively. The sorting of firms into no entry,

informal, formal, and exporter sectors based on their productivity draws is plotted in Figure G.19. As in Melitz (2003), the most productive firms enter the export market to sell at a higher international price.

Figure G.19: ZPC and free-entry with exports



### G.5.3 Equilibrium

To close the model, I need to specify the labor supply and the domestic product demand. Let  $L_i^{N,S}(w_i)$  and  $L_f^{N,S}(w_f)$  be the informal and formal labor supply curves of natives. In equilibrium, wages are determined such that the formal labor demand equals formal labor supply, and vice versa for the informal workers.

Unlike the baseline model, product market clearing no longer comes free. Let the domestic product demand be given by  $C(p)$ . Let  $q_s(\theta, p, w)$  denote the optimal production of firm with productivity  $\theta$  in sector  $s$  for given price  $p$  and wages  $w$ . In equilibrium, domestic product supply and demand determines the domestic price  $p$ .

$$\int_{\bar{\theta}_i}^{\bar{\theta}_f} q_i(\theta, p, w) dG(\theta) + \int_{\bar{\theta}_f}^{\bar{\theta}_x} q_f(\theta, p, w) dG(\theta) = C(p) \quad (22)$$

To summarize, in equilibrium (i) the zero profit cutoff and free entry conditions hold; (ii) labor markets clear, (iii) domestic product markets clear.

### G.5.4 Labor supply, product demand, and entrepreneurial effects of refugees

This model is rich enough to incorporate the empirical facts that refugees work, consume goods and services, and form businesses themselves. Let  $R$  denote the amount of refugees in the economy. Refugees' labor supply effect is captured by  $\frac{dL_i(w_i)}{dR} = \gamma$ , the same way as the baseline model. Refugees' product demand effect can be captured by an increase in the consumer base,  $\frac{dC(p)}{dR} = \phi$ . Lastly, the fact that refugees can form businesses is naturally captured by a change in the mass of potential entrepreneurs  $\frac{dM}{dR} = \psi$  and the distribution of productivity  $G(\cdot)$ . Quantifying these channels is outside of the scope of this paper. Instead, I provide some comparative statics on Table G.7. The effects of informal labor supply, product

Table G.7: Comparative Statistics

	$L_i^D$	$L_f^D$	$\bar{\theta}_i$	$\bar{\theta}_f$	$\bar{\theta}_e$
Informal Labor supply	0+	-	-	+	0
Product Demand	+	+	-	-	+
Entrepreneurial	+	+	0	0	0

demand, and entrepreneurial activity on Informal and Formal demand are straight-forward. Informal labor supply incentivizes formal firms to rely more heavily on informal labor instead of formal labor, which decreases formal labor demand. If the wage in the formal sector is fixed by a minimum wage, then the informal labor demand curve (by formal firms) remains unaffected. In contrast, increase in product demand incentivizes firms to produce more and hence hire more labor. Entrepreneurial effect basically increases the number of firms in the market, which increases labor demand. These opposing forces explain why I document a decline in native employment only in industries that employ refugees intensively. In services, the small labor supply effect is nullified by the demand and entrepreneurial effects. In manufacturing and agriculture, the labor supply effect is simply stronger.

Refugees' effect on firm entry is more nuanced. First, the labor supply effect decreases informal wages, which allows some very unproductive entrepreneurs to survive in the informal sector by reducing their costs of operation. Second, LS disproportionately benefits the informal firms, which incentivizes the marginally productive firms to become informal instead. As it is currently constructed, LS does not change the incentives of the marginal exporter firm. Mechanically, both the exporter and non-exporter types of firm with productivity  $\bar{\theta}_x$  hire the same amount of informal workers, hence they benefit the same amount. Most importantly, the marginally exporter firm is an inframarginal formal firm: it is too productive to consider the informal sector, hence there is no informalizing effect for these firms.

The product demand channel has clear effects on the least productive and most productive firms. It increases the profits of the informal and non-exporter formal firms. Consequently, it allows the least productive firms to survive in the informal sector, and induces the marginally exporter firm to sell domestically instead. Under reasonable assumptions (details are given in the online appendix), more purchasing power in the economy also leads to less informality. The price increases incentivizes firms to produce more, which is less costly for formal firms at the margin. Lastly, the entrepreneurial effects do not change the incentives of the marginal firms. They simply increase the mass of firms everywhere.

The 3 major effect combined explain why I document an increase in new firm formation

*only* for exporter firms, and not for non-exporter formal firms. Both the entrepreneurial and product demand effects increase the number of formal firms. This is why I find a strong correlation between population and the number of new firms in Figure [D.11a](#). More people means more demand and more entrepreneurs, which increase the number of new firms in this model. However, this effect is suppressed by the informalizing effects of an informal labor supply shock, leading to a null effect in the aggregate. This informalizing effect does not matter for the more productive exporter firms, for which entrepreneurial and demand effects oppose each other. In reality, exporter firms also sell domestically, which means the domestic product demand effect is less relevant empirically for the trade-off between becoming an exporter vs not. The entrepreneurial effect dominates and increases the number of new exporter firms.