ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 84/25

Formal Effects of Informal Labor: Evidence from Syrian Refugees in Turkey

Ahmet Gulek

www.rfberlin.com

OCTOBER 2025

Formal Effects of Informal Labor: Evidence from Syrian Refugees in Turkey

Authors

Ahmet Gulek

Reference

JEL Codes: D22, J15, J21, J46, J61

Keywords: Informality, Immigration, Refugee crises, Work permits

Recommended Citation: Ahmet Gulek (2025): Formal Effects of Informal Labor: Evidence from Syrian Refugees in Turkey. RFBerlin Discussion Paper No. 84/25

Access

Disclaimer

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

Formal Effects of Informal Labor
Evidence from Syrian Refugees in Turkey

Ahmet Gulek*

October 15, 2025

Abstract

I study the effects of Syrian refugees, who are denied work permits and thus can only work informally, on Turkish firms and workers. Using travel distance as an instrument for refugee location, I show that low-skill natives lose both informal and formal salaried jobs. I document two mechanisms: formal firms reduce their formal labor demand and new firms do not enter the formal economy. Estimates imply an elasticity of substitution of 10 between formal and informal workers. Counterfactual exercises predict that granting refugees work permits would have created up to 120,000 formal jobs in the economy through higher informal wages.

JEL Classification: D22, J15, J21, J46, J61.

Keywords: Informality, Immigration, Refugee crises, Work permits

*Gulek: Postdoctoral scholar at Stanford Immigration Policy Lab (e-mail: agulek@stanford.edu). I am grateful to Daron Acemoglu, Onur Altindag, Josh Angrist, Abhijit Banerjee, Amy Finkelstein, participants at the Society of Labor Economists Summer meeting, the MIT Labor Lunch and Development Lunch, seminar participants at Bilkent University and Koc 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, and the Institute for Humane Studies at George Mason University (grant no: IHS016850, IHS017975, IHS017988, IHS018799, IHS017343). An earlier version of this paper circulated in 2022 under the title "Formal Effects of Informal Labor Supply: Evidence from the Syrian Refugees in Turkey" (SSRN no: 4264865).

1 Introduction

The global refugee population has quadrupled in the last decade, from 11 million in 2012 to 46 million today (UNHCR, 2021). Two distinct characteristics separate this migration wave from previous patterns. First, most refugees are hosted by developing countries with substantial informal economies. Second, host countries frequently deny work permits to refugees to protect their native workers from refugee competition.[1] For example, Turkey hosts the world's largest refugee population, yet the vast majority of these refugees lack work permits. Consequently, the 3.6 million Syrian refugees constitute a massive informal labor supply shock, and their impacts depend on the dynamics between the informal and formal sectors. In this paper, I present empirical evidence documenting how firms and native workers respond to an informal labor supply shock, theoretical analysis explaining the relevant economic forces, and counterfactual exercises that quantify the labor market impacts of granting refugees work permits.

The Syrian refugee crisis in Turkey provides an ideal setting to examine how informal immigration affects both informal and formal sectors for several reasons. First, Turkey is a developing country where 40% of all employment is informal, allowing informal immigrants to participate in the labor market. Second, the overwhelming majority of Syrian refugees in Turkey lack work permits and must seek informal employment. Third, Syrians settled mainly in regions close to the border, creating quasi-experimental variation in immigrant intensity across local labor markets. Fourth, Turkish labor force surveys distinguish between formal and informal employment for natives, making it possible to investigate each sector separately.[2] Fifth, firm-level census data allow me to separate the intensive and extensive margin effects for the formal sector.

I first analyze the refugees' impact on natives' employment in salaried jobs. Identification comes from an exposure design, where the travel distance between Turkish and Syrian cities operates as an instrument for migrants' location choice. Adjusting for pre-trends that reflect regional convergence in Turkey, I find that low-skill Syrian immigration reduces salaried employment for low-skill natives while high-skill natives maintain their employment rates. Low-skill natives lose informal and formal salaried jobs despite virtually no Syrians working in the formal sector during this period. My estimates indicate that a 1 percentage point (pp) increase in the refugee/native ratio decreases native informal salaried employment by 0.17 pp and formal salaried employment by 0.13 pp among low-skill natives. 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. These effects are similar for men and women, concentrate in industries and firms with greater exposure to migrant labor supply, and are not driven by contemporaneous trade shocks.

Formal employment decreases along both intensive and extensive margins. The panel of formal firms from the census shows that small, informal-intensive incumbent firms reduce their formal

[1] The Turkish Minister of Work and Social Security explicitly articulated this concern: "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).

[2] 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: those with (without) social security are formal (informal) workers.

labor demand while large incumbents maintain theirs. More surprisingly, the number of new, low-productivity firms also declines despite increases in total population, electricity consumption (a proxy for total firm activity), and the number of new, high-productivity firms. This evidence indicates that informal Syrian immigration increases total firm activity while creating a missing mass of new formal firms, suggesting that marginal firms choose to remain unregistered to access informal labor more easily. Although Turkey lacks credible data on unregistered firms, these results strongly suggest that denying immigrants work permits also increases informality through firm creation decisions.

I further explore how the native workers respond to losing their salaried jobs. I find that immigrants *increase* men's non-salaried employment, primarily self-employment, and do not impact women's non-salaried employment. The distinction between salaried employment and self-employment is important because salaried jobs arise partly from firms' labor demand, while self-employment is solely a decision of labor supply. This result implies that the outside option for salaried positions is self-employment for men and home production for women. This is a novel finding in the immigration literature, which focuses on developed economies where self-employment constitutes a smaller portion of the labor markets.[3] These results suggest that developing countries exhibit different labor market adjustments to immigration specifically because self-employment offers a viable alternative to salaried work.

I rationalize my findings with an equilibrium model where firms utilize both the intensive and extensive margins of informality, as in Ulyssea (2018). Firms employ low-skill and high-skill labor for production. Whereas high-skill labor can only be hired formally (thus restricted to formal firms), low-skill labor can be hired both formally and informally. In this model, 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 given diminishing returns to labor. When informal and formal workers are gross substitutes in production, informal immigrants can incentivize firms to become more informal. This can happen both on the intensive margin, by formally registered firms replacing their formal employees with informal ones, and on the extensive margin, by marginal new firms remaining unregistered.

I use my empirical findings and data moments to estimate key model parameters. The results show an elasticity of substitution between formal and informal labor of approximately 10. This is one of the first papers to estimate this elasticity.[4] This high elasticity is consistent with the Turkish context, where informal and formal workers frequently occupy the same sectors and firms. This finding supports the assumption of perfect substitutability between informal and formal workers

[3] For a literature review, please refer to Dustmann et al. (2016).

[4] The only study that I could find is Schramm (2014), who studies the equilibrium effects of taxation on sectoral choice, work hours, and wages in Mexico. She finds an elasticity of 1.8, substantially lower than my estimate. This discrepancy likely stems from the fact that informal and formal workers occupy different sectors and firms in Mexico, whereas they work in the same firms in Turkey, which makes these two types of workers more substitutable.

used in recent structural literature on the informal sector (Ulyssea, 2018, 2020).

Finally, I use my model to estimate the labor market impacts of providing refugees with work permits. This counterfactual addresses critical policy concerns because (i) most refugees in the world do not have work permits (Clemens et al., 2018), and (ii) governments in both developing and developed countries recently started granting this right.[5] The model highlights a key trade-off for policymakers: work permits redirect some informal labor supply to the formal sector, which (i) increases wages and native employment in the informal sector through lessened competition and (ii) decreases native employment in the formal sector through increased competition. Higher informal wages also compel firms to seek more formal workers because of the high substitutability between the two factors. This indirect effect cannot overcome the direct competition effect. Thus, work permits necessarily decrease native formal employment. However, as firms demand more formal labor, work permits create more formal jobs in the economy. The model predicts that if refugees matched natives' formality rates, a 1 pp increase in refugee/native ratio would decrease informal salaried employment by 0.06 pp and formal salaried employment by 0.47 pp among natives. Despite refugees replacing more natives in the formal sector, work permits would create 120,000 additional formal jobs as firms substitute away from informal labor due to higher informal wages. As a benchmark, this would be equivalent to an 18% increase in GDP per capita in terms of creating formal jobs.[6]

This paper contributes to five strands of literature. First, it complements research studying the dynamics between the informal and formal sectors. Earlier contributions were largely theoretical (Rauch, 1991; Amaral and Quintin, 2006), while recent work has focused on estimating structural models (Bosch and Esteban-Pretel, 2012; Meghir et al., 2015; Ulyssea, 2018). I contribute to this literature by providing quasi-experimental evidence and theoretical analysis showing that informal immigration drives firms toward informality on both the intensive and extensive margins, indicating high substitutability between informal and formal labor in production. Similar claims on the intensive margin are made in contemporaneous work by Delgado-Prieto (2024), who shows that Venezuelan immigrants displaced Colombian natives in the formal sector but not in the informal sector, which he qualitatively attributes to the high substitutability between informal and formal workers. However, he does not quantify this elasticity because (1) Venezuelans supplied both informal and formal labor as they were granted work permits (Bahar et al., 2021), and (2) his modeling choices do not allow for estimation. His approach can thus be seen as complementary to my approach, which focuses on the intensive and extensive margin impacts of a fully informal labor supply shock, measures the substitutability between the factors, and quantifies the role that

[5] Examples include Colombia granting Venezuelan refugees work permits beginning in 2017 (Bahar et al., 2021), the US committing to provide permits to five hundred thousand Venezuelan refugees (Hesson, 2023), and Poland implementing early work permits for Ukrainian refugees (Lesinska, 2022).

[6] 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 salaried jobs decreased by 8 pp from 0.45 to 0.37. If 2004 informality rates persisted to 2011, 650,000 fewer formal jobs would exist. If economic growth alone caused this decrease in informality à la La Porta and Shleifer (2014), then providing work permits to refugees would be equivalent to an 18% growth in GDP per capita for creating formal jobs.

denying refugees work permits plays in these effects.

My counterfactual prediction on the formalization effects of work permits builds on 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) and developed countries (Elias et al., 2025). Two papers focusing on work permits in refugee crises are most relevant to my work. In policy work, Clemens et al. (2018) provide economic arguments explaining why refugee work permits can benefit both refugees and natives. Empirically, Bahar et al. (2021) study the effects of granting Venezuelan refugees work permits and find negative but negligible effects on the formal employment rate of Colombian workers.[7] My paper complements their findings by predicting that granting refugees work permits should create formal jobs in the economy. However, this comes at a cost to some natives who lose their formal jobs due to increased competition.

This paper builds on the extensive literature using refugee shocks to study the effects of immigration on labor markets. 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). I contribute to this literature in several ways. First, I study a massive shock. While the Mariel Boatlift increased Miami's adult population by 8%, Syrians increased Turkish cities' by up to 94%. Second, I show that labor market adjustments to immigration can be vastly different in developing and developed countries due to natives' ability to transition into non-salaried positions. Third, I provide a framework for thinking about the labor market consequences of modern refugee crises, where host countries have large informal sectors and governments can grant or withhold work permits from refugees.

More recently, several papers investigated the effects of Syrian refugees on 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). Using different identification strategies, they found inconclusive results. Del Carpio and Wagner (2015) find an increase in formal employment among only low-skill men. However, Akgündüz and Torun (2020) claim that high-skill employment (which is mostly formal) has increased. Among 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. Cengiz and Tekgüç (2022) claim that natives actually did not lose jobs due to the refugee shock. As Appendix Section F details, these contradictions stem from two factors: (i) misinterpreting differential trends as causal effects and (ii) overlooking men's transitions to non-salaried positions. My analysis adjusts for pre-trends, combines additional data sources, and separates salaried from non-salaried positions.

[7] I predict stronger disemployment of natives in the formal sector than Bahar et al. (2021). One potential explanation for our different conclusions is that I focus on salaried employment, whereas they study aggregate employment. If Colombian natives who lose their formal salaried jobs transition to formal non-salaried jobs, as I documented in Turkey, then our conclusions would be consistent.

This approach reveals that natives lost salaried jobs in both the informal and formal sectors. My theoretical framework explains these findings, identifies key economic mechanisms, and quantifies how withholding work permits generates these results.

Lastly, my results on firm entry complement a literature studying immigrants' effects on firm creation.[8] In developed countries, high-skill immigrants create businesses (Azoulay et al., 2022), and immigrant labor enables firm creation and faster startup growth (Dimmock et al., 2022). Evidence on the effects of low-skill immigration remains scarce, particularly in developing countries. Previous studies of Syrian refugees in Turkey document increases in foreign-owned firms (Akgündüz et al., 2018), firms with more than 20 employees (Altındağ et al., 2020), and firms participating in international trade (Akgündüz et al., 2023). These findings align with standard entry models: immigration increases population, and the market size effect enables new entrants (Seim, 2006). However, my analysis using more comprehensive data reveals a novel economic force. I document a missing mass of new, low-productivity firms in the formal sector, a surprising result that I show stems from immigrants lacking work permits. My findings demonstrate that denying refugees work permits incentivizes firms to become more informal through both intensive and extensive margins: formally registered firms reduce their formal labor demand, and marginal entrepreneurs choose to remain unregistered to access informal labor.

The rest of this paper is structured as follows. Section 2 introduces the conceptual framework and motivates the empirical analysis, Section 3 provides the necessary background on the Turkish labor markets and Syrian refugees in Turkey, Section 4 explains the identification strategy, Section 5 presents the empirical results, Section 6 introduces the model, and Section 7 concludes.

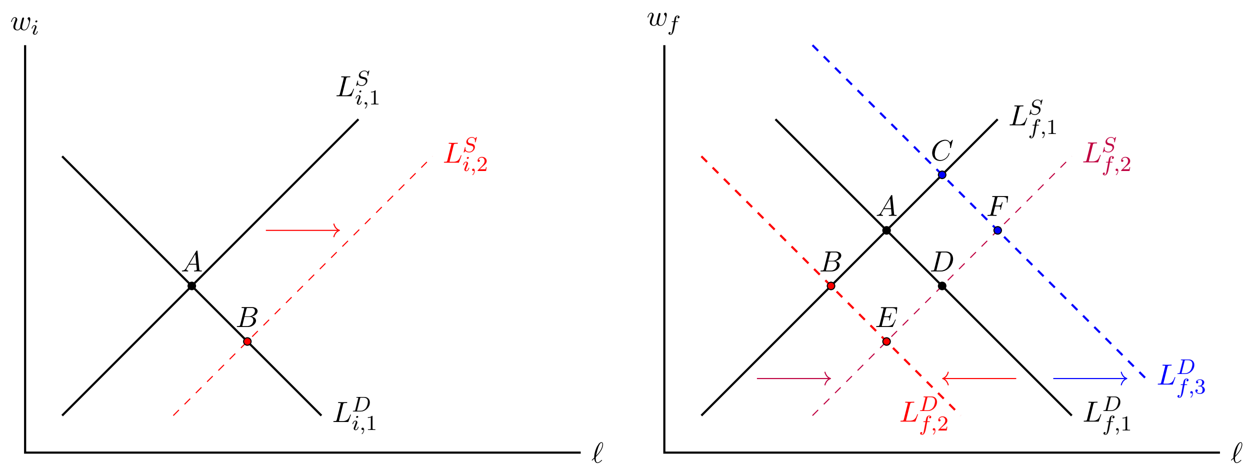
2 Conceptual Framework

This section motivates the empirical analysis by showing that the impacts of an informal labor supply shock on the formal sector cannot be signed by theory alone. Natives' employment rate in the formal sector can increase or decrease under different models of informality, which makes this a fundamentally empirical question. To build intuition, I visualize the potential equilibrium effects using labor supply and demand curves. I formalize these mechanisms in Section 6.

Figure 1 illustrates the potential changes in informal and formal labor markets upon the arrival of informal immigrant workers. First, under the canonical labor demand framework where firms utilize both informal and formal workers in production, an increase in informal labor supply necessarily reduces wages and displaces native employment in the informal sector. As shown in Figure 1a, this manifests as a rightward shift in the labor supply curve, causing the equilibrium to move from point A to B along the downward-sloping labor demand curve. While total informal employment in the economy increases, some native workers are displaced from the informal sector. How this decrease in informal wages subsequently affects the formal sector depends on the responses of both firms and native workers to this shock.

[8] For excellent reviews, see Lofstrom and Wang (2022); Chodavadia et al. (2024)

Figure 1: Effects of an informal labor supply on informal and formal sectors under different models



(a) Informal Sector

(b) Formal Sector

Notes: Panel A depicts the informal labor market with a rightward shift in labor supply following immigration. Panel B illustrates potential changes in the formal labor market under different theoretical scenarios. Points A, B, C, and D represent equilibria in both sectors. In the informal sector, informal immigrant arrival lowers wages and displaces some native workers (shift from A to B). This wage decrease affects the formal sector through changes in labor demand and supply curves. On the demand side, when informal and formal workers are substitutes, formal labor demand shifts leftward (equilibrium moves to B); when they are complements, formal labor demand shifts rightward (equilibrium moves to C). On the supply side, if native workers transition from the informal to the formal sector, formal labor supply shifts rightward. Depending on where the new labor demand curve resides, equilibrium can be any point within the BCEF rectangle.

Figure 1b illustrates the possible changes in formal sector equilibrium. The solid lines represent the initial labor demand and supply curves, with point A denoting the baseline equilibrium. The dashed lines indicate potential shifts in these curves following the informal labor supply shock, with points B, C, D, E and F representing alternative new equilibria under different economic forces.

The effects of lower informal wages on formal labor demand depend critically on the degree of substitutability between informal and formal workers in production. When these inputs are highly substitutable, as in Ulyssea (2018), a decrease in informal wages incentivizes firms to substitute toward informal and away from formal workers. This is depicted as a leftward shift in the formal labor demand curve, moving the equilibrium from point A to point B and resulting in employment and wage losses for native workers in the formal sector. In contrast, when informal and formal labor are gross complements, the decrease in informal wages induces firms to increase their demand for formal workers. This appears as a rightward shift in the formal labor demand curve, changing the equilibrium from point A to point C. Under this scenario, native workers gain employment opportunities and experience higher wages in the formal sector.

The effects of lower informal wages on formal labor supply depend on workers' ability to transition between sectors. In search models where native workers can endogenously sort between informal and formal sectors, as in Meghir et al. (2015), a decrease in informal wages causes natives to redirect their search efforts to the formal sector. Figure 1b represents this as a rightward shift in the formal labor supply curve. Absent changes in labor demand, this labor supply shift would move the equilibrium from point A to point D: natives obtain more formal jobs but observe wage losses.

Figure 1 demonstrates that the net effects on the formal sector cannot be determined by theory alone, making this a fundamentally empirical question.[9] Depending on the substitutability between informal and formal workers in production and workers' ability to reallocate their search efforts, any point within the BCEF rectangle becomes a potential equilibrium. The causal impacts of an informal labor supply shock on the formal sector reveal which mechanism dominates. For instance, substitutability between informal and formal workers reduces formal employment, whereas natives' ability to shift their labor supply to the formal sector increases it. Consequently, an observed decrease in formal employment would suggest that informal and formal labor are gross substitutes in production and that natives' capacity to transition to the formal sector is of second-order importance. My empirical analysis identifies which of these theoretical channels dominates in Turkey.

3 Data and Background

Native employment

Information about the informal and formal labor market outcomes of native workers comes from the 2004–2016 Turkish Household Labor Force Surveys (HLFS) conducted by the Turkish Statistical Institute (TurkSTAT). HLFS is representative at the NUTS-2 level, which consists of 26 regions. The sampling is based on the national address database and does not cover the Syrian refugees who are under temporary protection.

HLFS codes employment into four categories. Between 2004–2016, 61% of employed natives were regular salaried workers, 21% were self-employed, 13% were unpaid family workers, and 6% were employers. I combine the last three groups into one "non-salaried employment" category. This allows a tractable separation of jobs that are partly determined by the labor demand of firms and jobs that depend solely on individual labor supply decisions. This distinction is critical in studying how firms respond to informal labor supply shocks. For instance, consider a native who loses his formal, salaried job due to being replaced by informal refugees. This native may keep "working" as an unpaid family worker or trade items at the local markets as a self-employed person. Self-employment can also be formal if the worker pays his social security benefits. Either way, this native would appear as "employed" under the HLFS, even though his employer replaced him with informal immigrants. Consequently, focusing on the overall employment rate of natives misses how firms respond to an informal labor supply shock. To prevent this problem, I study salaried employment and non-salaried employment separately while focusing on salaried employment as

[9]The forces described here are not exhaustive. Additional effects could arise in models where informal and formal firms compete in the product market or rely on each other's goods as intermediates in production.

the key outcome of interest in both the theoretical and empirical analyses. The salaried and non-salaried employment statistics among different types of natives and industries can be found in Table A.1 in the Appendix.

I distinguish between formal and informal employment through workers' self-reported social insurance 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. Assuming that workers report truthfully in HLFS, I observe wages and employment status in both the formal and informal sectors. Although self-reported, insurance status is 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 informal sectors (Ulyssea, 2020). Across regions and industries, the informality rate (defined as the ratio of employment that is informal) decreases with education. It is higher in less developed regions and in industries like agriculture, construction, and textiles, which are known to rely on informal labor.

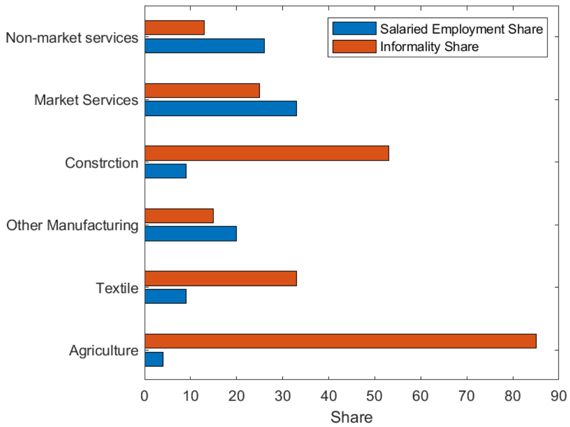
Figure 2 shows the informality rate across select industries and firm sizes. The informality rate is heterogeneous across sectors, ranging from 85% in agriculture to 13% in non-market services. Non-market services are mostly provided by the government, which explains the low informality rate. However, across all industries, informal and formal workers coexist. For example, in the textile industry, which has the highest proportion of refugee workers (Turkish Red Crescent and WFP, 2019), for every three salaried employees, one is informal and two are formal workers. Figure 2 shows the informality rate across firms of different sizes. Firms of all sizes rely on informal workers. The informality rate goes down drastically as firms get bigger: from 59% in firms with 1–9 employees to 29%, 16%, 7%, 4%, and 2% among firms with 10–24 employees, 25–49 employees, 50–249 employees, 250–499 employees, and more than 500 employees, respectively. This inverse relation between informality rate and firm size is well established in the literature and can be rationalized by larger firms being more visible and therefore having less room for illegal activities (Ulyssea, 2020).

I supplement the analysis on the formal sector by leveraging data from the Turkish census, which includes the universe of formal firms and workers from the tax records. The census excludes the informal sector but tracks formal firms over time, which provides two key advantages. First, it enables heterogeneity analysis by firm size to study effects across firms with varying informal intensities. Second, tracking incumbent firms before and after the shock isolates intensive margin adjustments.

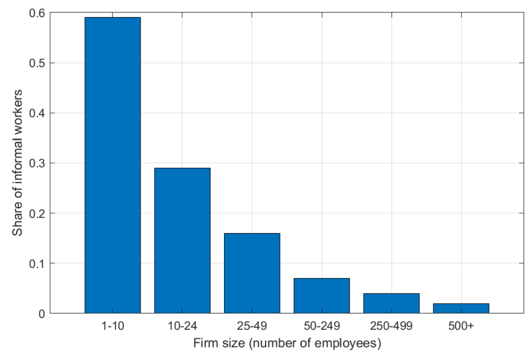
Firm entry

To study the extensive margin adjustment of firms, i.e., firms' decision to register with tax authorities, I leverage data on firm formation from three different sources. First, the Union of Chambers and Commodity Exchanges of Turkey (TOBB in Turkish) publishes the number of

Figure 2: Ratio of informal workers across industries and firm size



(a) Industries



(b) Firm size

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 (registered) 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 and more susceptible to extensive margin informality. 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 participating in international trade. In an average year, there are 109 thousand new incorporated firms in Turkey. The average number of new formal firms (including sole proprietorships) is around 350 thousand in Turkstat and 304 thousand at GBS.[10] Of these firms in GBS, 8.7 thousand export and 9.1 thousand import at least once in their lifetime.

Turkish institutions do not collect data on informal/unregistered firms. Therefore, I do not have a good estimate of the ratio of new firms that remain unregistered. Ozar (2003) is the only rigorous data collection effort on informal firms. She finds that around 4% of firms self-declare that they are not registered. The actual number is likely higher because, unlike working informally, operating an unregistered business is a crime. Consequently, informal firms have incentives to either not be interviewed or lie during interviews.[11] Moreover, 4% of firms being informal is an equilibrium outcome. If new informal firms have higher exit probabilities than new formal firms, then the ratio of informal firms among new firms would be higher. For example, Ulyssea (2018) estimates that the exit probability of unregistered firms is three times that of formal firms in Brazil. If this ratio

[10] Turktstat and GBS data do not match exactly, which is due to the different administrative sources they draw the data from. However, my qualitative results remain robust when using either data source.

[11] 4% firm informality is arguably too low for a country with 40% labor informality. As a comparison, Turkey and Brazil had similar GDP per capita and labor informality (40% and 46%, respectively) in 2011. Yet, 30% of firms with less than five employees in Brazil are unregistered (Ulyssea, 2018).

is similar in Turkey, this would imply that at least 12% of the new firms in Turkey in a given year are informal.

Additional Data sources

I utilize various data sources for robustness checks. Region-by-country level foreign trade statistics were gathered from Turkstat's Foreign Trade Statistics Micro Data Set. I use this data to study the trade shocks stemming from the Syrian War in the Appendix Section A. Moreover, I also utilize provincial electricity consumption data from Turkstat as a proxy for total (formal and informal) firm activity.

Lastly, the number of refugees in Turkey across years and provinces are acquired from the Directorate General of Migration Management of Turkey (DGMM). I use this data to determine the treatment intensity across years and regions. Unfortunately, DGMM does not share the education and age break-down of refugees at the province level, which prevents the empirical investigation from exploiting that variation.

Syrian Refugee Crisis in Turkey

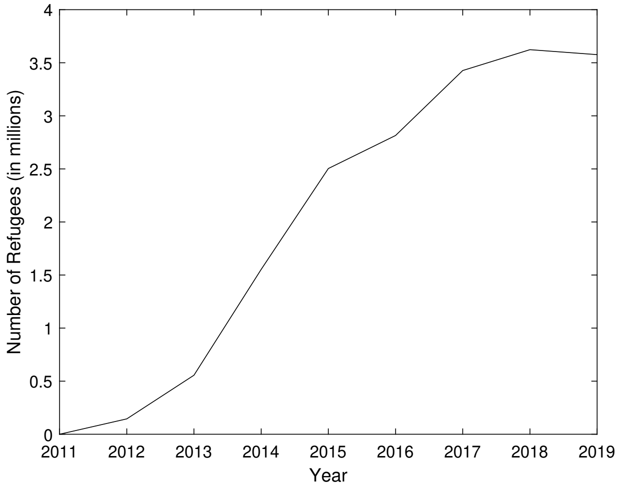
The Syrian Civil War began in March 2011. By 2017, 6 million Syrians had sought shelter outside of Syria, primarily in the neighboring countries Turkey, Lebanon, Jordan, and Iraq. With 3.6 million registered Syrian refugees, Turkey hosts the highest number of refugees in the world (UNHCR, 2022). The first waves of refugees began arriving in Turkey in late 2011, but their 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 in Turkey. Figure 3a shows how the number of Syrian refugees in Turkey has evolved. There were around 170 thousand refugees by 2012, 500 thousand by 2013, 1.6 million by 2014, 2.5 million by 2015, and nearly 3 million by 2016.

Syrian refugees are disproportionately less educated than Turkish natives. Figure 3b compares the education levels of Syrian refugees in Turkey with those 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 61% of Turkish natives. Taking into account the potential educational downgrading (Dustmann et al., 2013) and the fact that most Syrian refugees have only basic Turkish language skills (Turkish Red Crescent and WFP, 2019), the Syrian refugee shock can be interpreted as a low-skill labor supply shock for the Turkish labor markets.

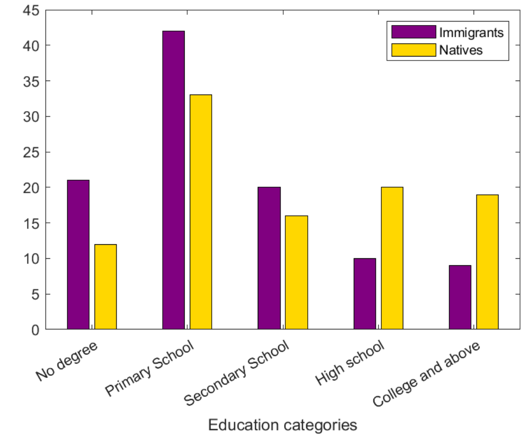
The Turkish government initially tried to host the Syrians in refugee camps in the southeastern part of the country across the Turkish-Syrian border. However, these camps quickly exceeded capacity as the number of arriving refugees increased. The refugees thus dispersed across Turkey.[12] Figure 3c shows the distribution of the number of Syrian refugees per 100 natives in Turkey at

[12] By 2017, only 8% of the refugees lived inside the camps.

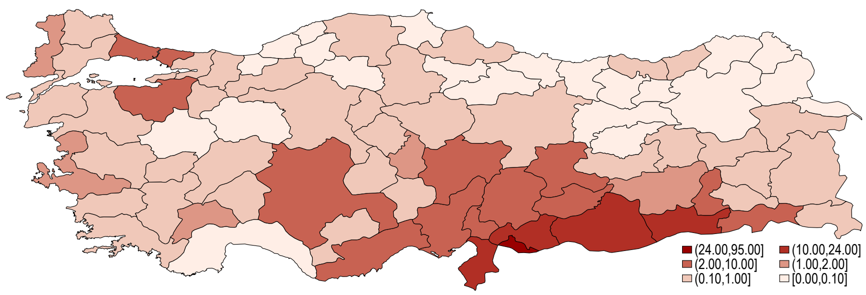
Figure 3: Statistics on the Syrian refugees in Turkey



(a) Timeseries of the number of Syrian refugees in Turkey



(b) Educational Attainment of Syrians and Natives



(c) Share of Syrian refugees in Turkish population (in%) in 2015

the province level. Refugees are more densely located in regions closer to the border. Distance to the populous governorates in Syria strongly predicts the number of refugees per native in a given region, which constitutes the backbone of my identification strategy.

Most Syrians came under the temporary protection category, which allows access to health care and education and allows freedom of movement.[13] Since the temporary protection regime does not offer work permits, the vast majority of the Syrian labor force works in the informal sector. By the end of 2015, only around 7,300 work permits were issued for 2.5 million Syrian refugees residing in Turkey (Aslan, 2016).

Labor force surveys conducted by the Turkish Statistical Institute do not sample refugees. To understand the employment outcomes of Syrian refugees, I use data from randomized surveys conducted on ESSN applicants by the Turkish Red Crescent and WFP. ESSN applicants are a selected sample, and WFP's questions on labor market activity differ from those in HLFS. This complicates the interpretability of these estimates. Nonetheless, they shed some light on how

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

refugees may have impacted the Turkish labor market.

According to these surveys, refugees have an astonishing 84% employment rate compared 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. Refugees have a comparative disadvantage in industries requiring language skills, since only 3% are proficient in Turkish. Perhaps not surprisingly, refugees work primarily in textiles (19%), construction (12%), and agriculture (10%). 47% of employed refugees work in regular jobs, defined as a job with a fixed salary and working hours. This is more restrictive than the salaried employment definition used by Turkstat, so the salaried employment rate of refugees should be even higher.[14] Textiles also have the highest share of refugees in regular positions, as 79% of the workers have regular positions. The average monthly income of refugees was 1058 TRY in 2019. In contrast, natives in the informal sector made 1565 TRY per month on average in the same year.

4 Identification

The identification strategy exploits the differential intensity of Syrian refugees across region-year cells. The treatment $R_{p,t}$ denotes the number of refugees per native in region p and year t. The key outcomes of interest are natives' salaried employment rates in the informal and formal sectors. If the local labor market conditions impact refugee settlement, then a simple difference in differences strategy would give biased estimates.

To circumvent this bias, I exploit the fact that travel distance strongly predicts migrant settlement in forced migration episodes (Angrist and Kugler, 2003; Del Carpio and Wagner, 2015). The weighted-distance instrument Z_p calculates the inverse travel distance between each Turkish region p and Syrian governorate s and takes an average using weights λ_s,

$$Z_p = \sum_{s=1}^{14} \lambda_s \frac{1}{d_{p,s}} \tag{1}$$

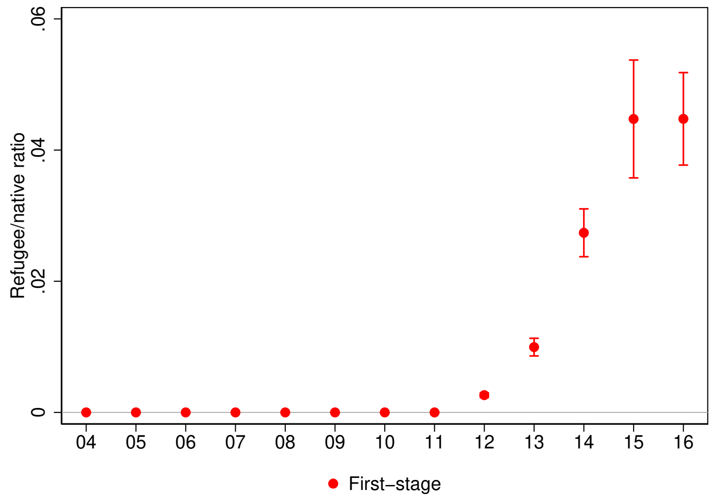
where $d_{p,s}$ is the travel distance between Turkish region p and Syrian governorate s, and λ_s is the weight given to Syrian governorate s.[15] Different weights λ have been used in the literature. In practice, weights matter little. I use the weights suggested by Aksu et al. (2022),

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

[14] For example, most work in construction is salaried but irregular.

[15] I use city centers in each region to calculate the travel distance. The data is available upon request.

Figure 4: Event study estimates of the first-stage



Notes: The regression equation is: $R_{p,t} = \sum_{j \neq 2010} \theta_j(\text{year}_j \times Z_p) + f_p + f_t + \eta_{p,t}$, where the instrument Z_p is standardized 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.

where $d_{s,c} \quad c \in \{T, L, J, I\}$ is the travel distance between Syrian region s to Turkey, Lebanon, Jordan, and Iraq respectively; and π_s is the population share in 2011, which I calculate using the 2011 census undertaken by the Central Bureau of Statistics of Syria. These weights predict how many refugees come from a Syrian region based on its population and proximity to Turkey compared to other bordering countries.

I use the instrument Z_p within both nonparametric and parametric event study models.

Nonparametric Event Study

The primary advantage of the nonparametric design is that it allows me to visually and flexibly assess the pattern of outcomes the distance instrument captures relative to the beginning of the refugee crisis. The basic nonparametric event study specification takes the form

$$y_{p,t} = \sum_{j \neq 2010} \theta_j(\text{year}_j \times Z_p) + f_p + f_t + \epsilon_{p,t} \tag{3}$$

where the instrument Z_p is standardized to have economically meaningful coefficients; f_p and f_t are region and year fixed effects. The standard errors are clustered at the region level. Figure 4 displays the estimates of θ_j from the first stage regression. Since there are no refugees in Turkey before 2012, $\theta_j = 0$ if $j < 2012$. The instrument strongly predicts refugee settlement in all post-treatment periods. The instrument's joint F-statistic in the years 2012–2016 is 238.

The figure reveals an increase in instrument-predicted treatment intensity over time. The intensity remained low in 2012 and 2013 due to fewer refugees and grew substantially afterward.

This time-series variation provides a visual check of the identification strategy: any causal effect of refugees should also increase over time.

The identifying assumption in this exposure design is that the instrument is orthogonal to local economic trends. However, this does not hold for several of the outcomes in the current setting. During 2004–2010 (before the refugee shock began), regions near the border observed higher growth in employment rates and wages, leading to a positive trend that is correlated with the instrument.[16]

To make progress, I exploit the empirical fact that pre-trends are approximately linear for most of the outcomes of interest throughout the paper. This guides my formulation of the parametric event studies that deliver the main estimates.

Parametric Event Study

I use the parametric event study to summarize the magnitude of estimated reduced-form effects and their statistical significance. The estimating equation and the presentation of results follow Dobkin et al. (2018) very closely. My choice of the functional form is guided by the patterns seen in the nonparametric event studies. In the figures below, I superimpose the estimated parametric event study on the nonparametric event study coefficients, which allows for a visual assessment of my parametric assumptions. In particular, the baseline specification is

$$y_{p,t} = \sum_{j \geq 2011} \beta_j (\text{year}_j \times Z_p) + \gamma Z_p t + \delta_p + \delta_t + \epsilon_{p,t} \tag{4}$$

Equation 4 includes a linear trend in instrument exposure $Z_p t$, which allows for regions to follow different trends that are correlated with their distance exposure. The key coefficients of interest, the β_js, show the change in the outcome predicted by the instrument relative to any pre-existing linear trend γ. As before, I include region and time dummies in the regression.

Interpretation

The parametric event study allows for a linear trend by distance exposure. The choice of the linear trend is motivated by the results from the nonparametric event studies, which, as we will see in the results below, suggest that a linear trend captures the differences in regional trends quite well. For the parametric event study, the identification assumption is that the distance to the border is orthogonal to *deviations* from the linear trend.

Accounting for pre-trends is one of the reasons why this paper documents novel empirical results. Appendix Section F presents a thorough discussion of the shortcomings of the identification strategies used in this literature. In short, no other strategy adequately addresses the fact that the border regions were catching up to the rest of the Turkish economy before the refugee crisis began.

[16] These pre-trends can be seen in the event study figures in the Appendix Section G.2.

IV Design

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

$$\begin{aligned} y_{p,t} &= \beta R_{p,t} + \delta Z_p t + f_p + f_t + \epsilon_{p,t} \\ R_{p,t} &= \sum_{j \geq 2011} \theta_j (\text{year}_j \times Z_p) + \gamma Z_p t + g_p + g_t + \eta_{p,t} \end{aligned} \tag{5}$$

where the treatment $R_{p,t}$ is instrumented by the interaction of distance Z_p with year dummies in the post-period; δ and γ are the linear trends in the structural and first-stage equations, respectively. Instrumenting the treatment R with a full set of interactions of distance and post-year dummies ensures that the linear trend is estimated using only the pre-period variation in both equations.[17]

Threats to Identification

Several potential threats to identification merit discussion. The distance instrument compares regions close to the border with those further away, and this comparison may fail to identify the causal effect of refugees for two main reasons.

First, the empirical strategy assumes that the Syrian war's impact on Turkish local labor markets should be orthogonal to distance from the border. This assumption could fail if Syria had been a major trade partner of border regions and the war had significantly disrupted trade flows. However, Syria was not a major trade partner of any Turkish region. Moreover, Appendix Figure C.4 shows that although trade initially fell in 2011 and 2012 at the war's onset, it more than recovered in border regions after 2013. Hence, no significant trade shock could have impacted local labor markets.

Second, the identification strategy assumes that regions constitute separate local labor markets and that immigrants' arrival in host regions does not affect labor markets in non-host regions. This imposes the stable unit treatment value assumption (SUTVA). Violation of this assumption would cause me to estimate *relative* effects rather than the *total* effects of immigration (Dustmann et al., 2016). Markets could re-equilibrate across space through population movements (Monras, 2020) or trade adjustments (Gulek and Garg, 2025). However, evidence suggests that such adjustments are not empirically relevant in this setting. Regarding labor supply spillovers, only minor changes occurred in population movements across space before 2016. Figure C.3 shows that regions closer to the border experienced slightly more out-migration and less in-migration, but these effects are too small to bias the IV estimates in an economically meaningful way. Regarding trade spillovers, Gulek and Garg (2025) demonstrate that the impact on labor demand in non-host regions has been minimal. The authors document that while immigration to major cities like Istanbul and Ankara would have generated large general equilibrium effects, refugees are located mostly in border regions

[17] This technical detail turns out to be pivotal in addressing the correlation between the instrument and the regional trends. More details can be found in Appendix Section F.

that are not central nodes of the domestic trade network and, therefore, generate only negligible spillovers. Consequently, potential SUTVA violations are not a first-order concern.[18]

5 Empirical Results

This section presents empirical estimates in four parts. First, it demonstrates that Syrian immigration caused low-skill natives to lose both informal and formal salaried positions. Second, it summarizes evidence that eliminates several potential confounders. Third, it documents native men's transition into non-salaried positions after losing salaried jobs. Fourth, it reveals changes in firm entry margins, which indicate that marginal firms are choosing to remain unregistered.

5.1 Low-skill Natives Lose Informal and Formal Jobs

This section shows the effects of refugees on natives' labor market outcomes. It focuses on the impact on salaried employment to capture the changes in labor demand. Since Syrian refugees in Turkey supply predominantly low-skilled labor, I separately analyze natives without a high-school degree (low-skill) and with at least a high-school degree (high-skill). This analysis reveals that low-skill natives lose salaried jobs while high-skill natives do not. To see where the low-skill employment losses are coming from, I separately analyze the informal and formal employment rates, which shows that low-skill natives lose both informal and formal jobs. This highlights that informal and formal workers are largely substitutable.

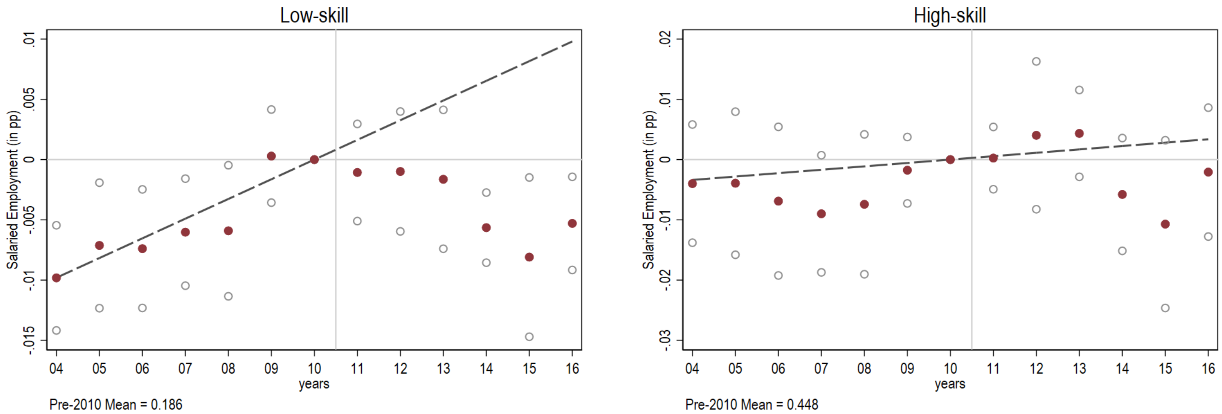
It is important to note that I focus on salaried employment rather than wages because the survey data is not a panel of individuals. Without the ability to track people over time, I cannot account for compositional changes. For example, if natives who lose their jobs are the lowest earners, the average wage conditional on working would increase even when no worker observes a wage increase.

Event Study Estimates

I begin by estimating the nonparametric and parametric event study designs shown in equations 3 and 4. Figure 5 plots the point estimates from the nonparametric design and the linear trend from the parametric design. Figure 5a shows the results on low-skill natives' salaried employment rates. There are two important results. First, there is a significant pre-trend: from 2004 to 2010 (before the treatment), regions closer to the border observed larger increases in salaried employment for low-skill natives. Notice that this trend was highly linear. The linear trend estimated in the parametric design not only falls under the 95% confidence intervals of the nonparametric estimates in the pre-period but is also very close to the point estimates. Second, the estimated effects from the parametric design, which are the differences between the nonparametric estimates and the linear

[18] The analysis ends in 2016 for several reasons, including a minimum wage increase and the beginning of the Emergency Social Safety Net (ESSN) program in which refugees received relatively large cash transfers. Both confounders could make it difficult to interpret the estimated effects post-2016.

Figure 5: Refugees' effects on native salaried employment



(a) Low-skill

(b) High-skill

Notes: The points in each figure represent the estimated effects of event time shown in equation 3. The hollow circles present the 95 percent confidence intervals. The dashed line represents the estimated pre-2010 linear relationship between outcome and instrument * event time from the parametric event study in equation 4 with the level normalized to match the nonparametric estimates.

trend, increase after 2013 in line with the refugee shock. The estimated effect is negative, meaning that Syrian immigrants caused low-skill natives to lose salaried jobs.

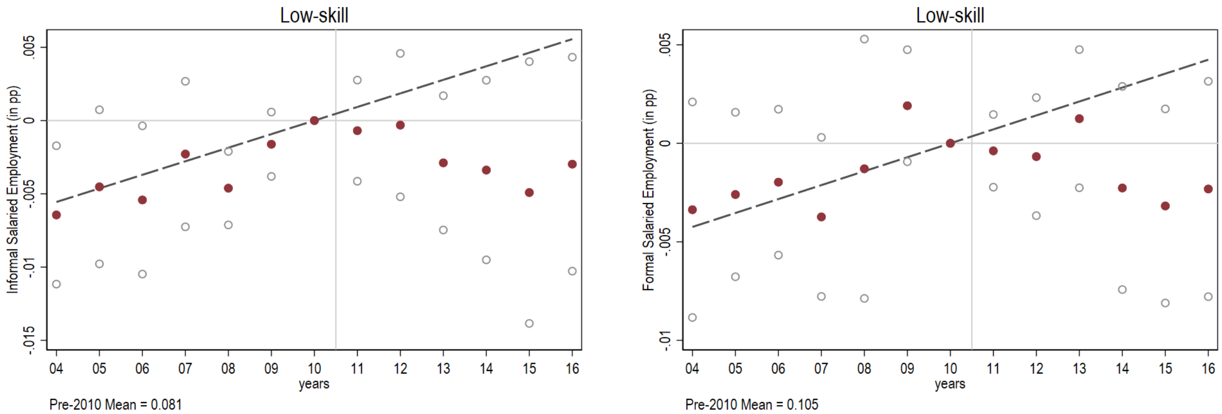
Figure 5b plots the results on high-skill natives. I document no economically meaningful pre-trend and no statistically significant deviation from the trend. I conclude that the Syrian immigrants displaced low-skill natives and did not impact high-skill natives' employment probabilities in the aggregate. This is intuitive as low-skill Syrians are a closer substitute for low-skill natives in the labor force and hence replace predominantly low-skill natives. The complementarity between low-skill and high-skill labor offsets the increased competition effect in the labor force, leading to a null impact on high-skill natives.

I continue by analyzing the informal and formal sectors separately. Figure 6 shows the results. First, I document economically meaningful pre-trends in both the informal and formal salaried employment rates of low-skill natives. These trends appear approximately linear, similar to the trend in Figure 5a, which increases the expected validity of the identification strategy. Lastly, I find statistically significant decreases (compared to the trend) in both the informal and formal sectors in the post-period. As the conceptual framework shows, these results imply that informal and formal workers are highly substitutable in production.

2SLS Estimates

To get economically meaningful estimates, I estimate equation 5 using 2SLS. The first row of Figure 7 shows the estimated effects of refugees on the informal and formal salaried employment of natives. A 1 pp increase in the the refugee/native ratio decreases the informal salaried employment

Figure 6: Informal/Formal composition of refugees' effects on low-skill natives



(a) Informal

(b) Formal

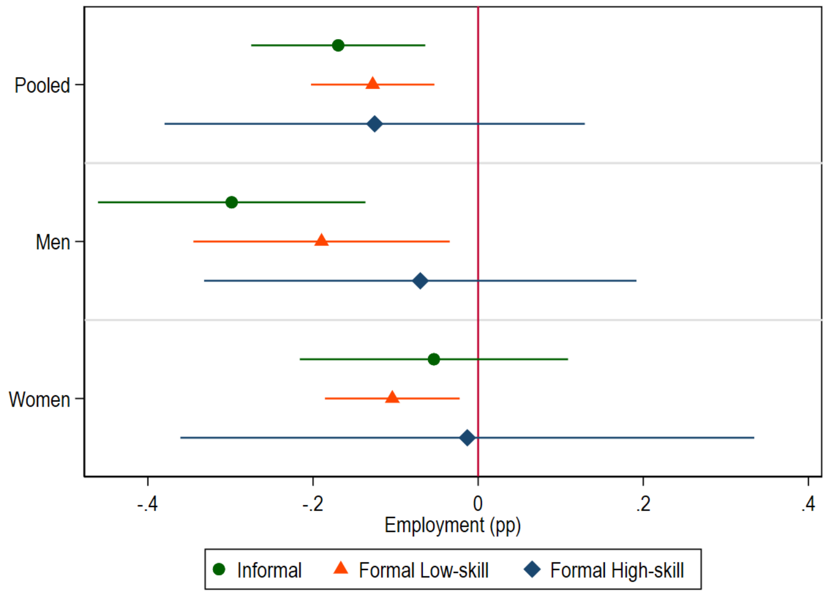
Notes: The points in each figure represent the estimated effects of event time shown in equation 3. The hollow circles present the 95 percent confidence intervals. The dashed line represents the estimated pre-2010 linear relationship between outcome and instrument * event time from the parametric event study in equation 4 with the level normalized to match the nonparametric estimates.

rate of natives by 0.17 pp, the formal salaried employment rate of low-skill natives by 0.13 pp, and does not significantly impact the formal salaried employment rate of high-skill natives in the aggregate. The second and third rows of Figure 7 separate these effects by sex. A 1 pp increase in refugee/native ratio decreases men's informal salaried employment rate by 0.30 pp and low-skill formal employment by 0.19 pp. For women, these effects are 0.05 and 0.10, respectively, with only the effect on formal employment being statistically significant. Lastly, there are no significant effects on the formal salaried employment rates of high-skill men and women.

While the immigration shock replaces some natives, it increases the total number of workers in the economy. 47% of ESSN applicants were working in regular jobs with fixed salaries and working hours in 2019 (Turkish Red Crescent and WFP, 2019). This is more restrictive than the salaried employment definition used by the TurkSTAT, so the salaried 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 in Turkey, 45 were working as salaried workers. 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 salaried workers, meaning 233 salaried natives. In period 2, this region receives 100 refugees, a 10 pp increase in refugee/native ratio. My estimates suggest that this shock leads to 30 natives losing informal and formal salaried jobs combined. In other words, 45 working refugees replace 30 natives. The total low-skill employment increases by $15/233 = 6.4\%$.

These estimates suggest that the informal refugee shock has caused native disemployment in both the informal and formal sectors. My preferred interpretation is that an informal labor supply shock incentivizes firms to become more informal by replacing their formal (and informal) native

Figure 7: Refugees' effects on native salaried employment rates



Notes: The 2SLS estimates come from estimating equation 5 using natives' informal, low-skill formal, and high-skill formal salaried employment rates. The first row shows the estimates using the pooled data. The second and third rows condition the sample on men and women separately. The 95% confidence intervals are plotted.

workers with informal refugees. However, there are alternative mechanisms that could create native disemployment in the formal sector. In a model where only unregistered/informal firms can employ informal workers and informal and formal firms compete in the product market, an informal labor supply shock would cause formal firms to shrink due to business stealing. This would reduce formal labor demand and create native disemployment in the formal sector. Alternatively, refugees demanding mostly the goods and services of informal firms could also reduce formal labor demand in general equilibrium. However, I can rule out these demand side channels because only low-skill natives lose jobs in the formal sector. This is consistent with (low-skill) refugees being closer substitutes in production to low-skill natives but inconsistent with these alternative models. The evidence suggests that formal firms can substitute between formal and informal workers among the low-skilled. Before further exploring the implications of these findings, I investigate their robustness.

5.2 Supporting Evidence

Effects Concentrate in Immigrant-intensive Industries

Syrian refugees disproportionately work in particular industries due to comparative advantage. Most are not proficient in Turkish, making them less likely to perform tasks that require written or spoken communication. Consequently, they predominantly work in jobs that require manual work: textiles (19%), construction (12%), and agriculture (8%) (Turkish Red Crescent and WFP, 2019). Appendix Figure B.1 shows that native job losses come primarily from these refugee-hiring industries. In fact, natives lost most of their jobs in textiles. This implies that it is the labor supply shock from refugees, not an unobserved negative demand shock, that causes low-skill natives to lose jobs.

Effects Concentrate in Informal-intensive Small Firms

Small firms are more informal labor intensive in Turkey. Consequently, if the native disemployment is due to the informal labor supply shock, then the decrease in formal employment should concentrate in small firms. To test this hypothesis, I use census data to group textile firms, the group with the largest native employment losses, into two categories: those with less than 50 formal employees in 2010 and those with 50 or more employees. I then estimate the effects of refugees on (the natural logarithm of) the number of employees of firms separately on small and large textile firms. Appendix Figure B.2 shows that, despite small and large firms following similar linear trends in the pre-period, small firms deviate from their trend and shrink, while large firms continue their trend. This evidence not only suggests that the labor supply of immigrants caused natives to lose salaried jobs, but it also provides further validation for the linear trend assumption. The less-exposed large firms continue their trend in the post-period.

Results Remain Robust to Nonparametric Pre-trend Adjustments

My identification strategy relies on the assumption that distance from the border is orthogonal to deviations from the linear trend. This could fail, for example, if the convergence between the southeast and northwest regions slowed down in the post-period. To account for potential deviations from the trend, I employ the Synthetic instrumental variable (SIV) algorithm (Gulek and Vives-i Bastida, 2025). SIV applies Synthetic Controls to account for pre-trends while still relying on the weights assigned by the instruments for identification. Appendix Section C.3 provides the details of the implementation, and Figure C.5 replicates the main results. My main conclusion remains robust: the informal labor supply shock causes natives to lose both informal and formal salaried jobs.

Results Help Reconcile the Contradictory Estimates in the Literature

My claim that the Syrian refugees caused low-skill natives to lose informal and formal jobs contradicts some of the existing evidence in published papers, as summarized in the Introduction. Given that this represents the world's largest refugee crisis, establishing reliable causal estimates and understanding the underlying economic mechanisms is of substantial scientific and policy importance. Appendix Section F provides a detailed comparison of my findings with existing studies and examines methodological issues in several published papers. The analysis reveals that previous studies suffer from one or more of three critical limitations: failure to adequately adjust for pre-existing regional trends, inadequate treatment of refugee location endogeneity, and overlooking that salaried and non-salaried jobs are driven by different economic forces, as I show in the next section. Each of these methodological shortcomings generates bias that fundamentally alters the interpretation of the economic forces at play, thereby explaining the literature's conflicting conclusions.

5.3 Natives' Escape to Non-salaried Jobs

Until now, I have focused on salaried employment instead of overall employment. This subsection shows the importance of separating salaried employment from non-salaried employment when studying the labor demand responses to refugee inflows. This distinction helps explain why this paper's empirical findings diverge from the existing literature on the labor market consequences of Syrian refugees in Turkey.

As Section 3 details, there exists an economically significant distinction between salaried and non-salaried employment in Turkey, a pattern generalizable to similar developing economies. Salaried employment is characterized by jobs where a worker's employment status is contingent upon an employer's decision. If an employer identifies less expensive labor to perform equivalent tasks, the worker risks displacement. In contrast, individuals engaged in any market activity, regardless of scale or formality, can accurately classify themselves as self-employed. For instance, when refugees displace natives from salaried positions in the textile industry, displaced natives with strong labor force attachment may maintain their *employed* status by engaging in alternative market activities independently. Consequently, the net effect on aggregate employment statistics may appear negligible despite substantial displacement from salaried positions. To empirically demonstrate this mechanism, I separately estimate the effect of refugee inflows on natives' total, salaried, and non-salaried employment rates within the formal sector, focusing specifically on low-skilled natives.

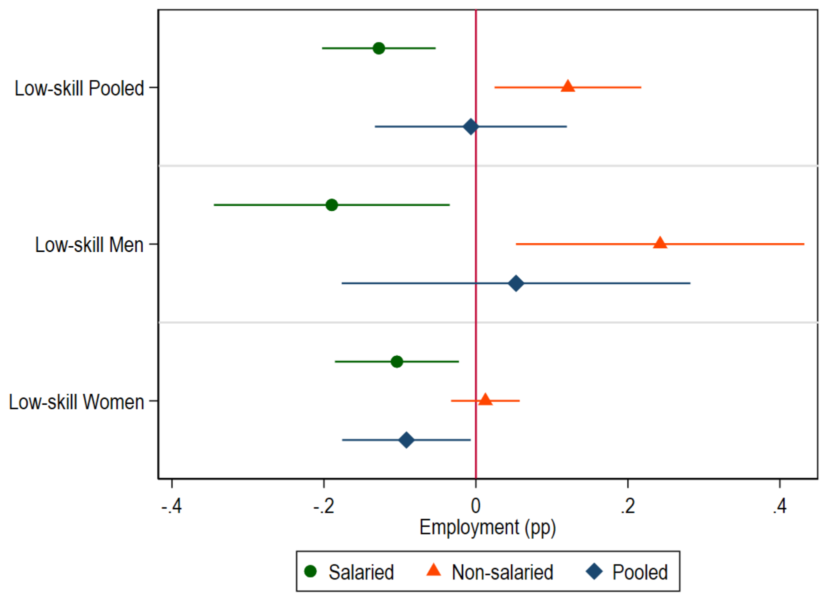
Figure 8 presents the estimation results for these employment outcomes.[19] Examining the first row, I find no statistically significant effect of refugee inflows on natives' total employment rates. However, this aggregate null effect conceals substantial heterogeneity across employment types. Consistent with the findings in previous sections, natives' formal salaried employment rates decrease significantly in response to refugee inflows. Importantly, this decline is counterbalanced by a nearly equivalent increase in non-salaried employment rates.

The second and third rows reveal substantial heterogeneity across men and women. Whereas they both experience similar decreases in salaried employment, only men transition into non-salaried jobs. This shows that the outside option of losing salaried jobs is self-employment for men and leisure or home production for women.[20] Further supporting this interpretation, Appendix Figure G.2 documents pronounced heterogeneity across industries. The decline in formal salaried employment predominantly affects the textile industry, while the rise in non-salaried work is largely

[19]The corresponding event-study estimates are available in Figure G.2 in Appendix Section G.2.

[20]This gender heterogeneity likely reflects socioeconomic norms in Turkish households, where men typically serve as primary income earners. When displaced from salaried positions, men may face stronger economic incentives and social expectations to maintain income-generating activities through self-employment while searching for new salaried opportunities.

Figure 8: Refugees' effects on salaried and non-salaried employment rates of low-skill natives



Notes: The 2SLS estimates come from the IV desing in equation 5. The first row shows the estimates using the pooled data on low-skilled natives. The second and third rows condition the sample on men and women separately. Standard errors are clustered at the nuts2 region level. The 95% confidence intervals are plotted.

seen in the services sector. This observation aligns intuitively with the opportunities available to self-employed individuals. It is much harder for a laid-off textile worker to open a textile shop than to buy and sell goods in the market.

An alternative explanation warrants consideration: refugee inflows might increase demand in non-tradable services, creating preferable job opportunities that natives voluntarily choose over textile sector employment. However, three key patterns in the data contradict this hypothesis. First, the observed employment gains are exclusively in non-salaried positions, with no corresponding increase in salaried service sector employment. Second, a demand-driven explanation would predict employment shifts across multiple industries, yet formal salaried employment remains stable in sectors without refugee workers. Third, this alternative cannot explain the gender asymmetry: why both men and women lose salaried jobs, but only men transition to non-salaried employment. The evidence consistently supports the displacement of formally employed natives by informal refugee workers rather than voluntary job transitions driven by demand effects.

Figure 8 illustrates why analyses focused on overall employment have led to misinterpretations in prior literature. For example, Cengiz and Tekgüç (2022) concluded that Syrian refugees did not reduce native employment, while Erten and Keskin (2021) argued that Syrians displaced Turkish women but not men. These divergent conclusions stem from overlooking the critical distinction between employment types. When self-employment serves as an outside option to salaried work, aggregate employment effects become misleading indicators of displacement. The results demonstrate that both men and women lose salaried jobs, but men's almost one-to-one transition to self-employment masks the true displacement effects, underscoring the importance of disaggregating employment categories when studying labor market impacts of immigration.[21]

All of the estimates shown 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 G.1, G.2, and G.3 in the Appendix Section G. The results are robust across different cuts of the data.

5.4 Firms' Escape to Informal Sector

Native displacement from formal employment can occur through two distinct mechanisms. First, on the intensive margin, formally registered firms may substitute formal native employees with informal refugee workers. Second, on the extensive margin, new enterprises that would typically enter the formal sector may instead remain unregistered, limiting their demand to informal labor. The previous section documented evidence for the intensive margin through reduced formal labor demand among small textile firms. This section examines the extensive margin mechanism.

The identification challenge in this section is more nuanced. First, refugees increase the local population immensely and, therefore, can increase the formation of new firms (Seim, 2006). In contrast, if there are marginal entrepreneurs who 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 informal firm entry is not observed. Therefore, these two channels cannot be estimated separately.

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

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

I first estimate the nonparametric event study design shown in equation 3, where the outcome

[21] Notably, Aksu et al. (2022) also investigate heterogeneity across different employment types, but the bias inherent in their empirical design, as detailed in Appendix Section F, results in them wrongly claiming that women also transition from salaried to non-salaried positions.

variable is the natural logarithm of the number of (i) all firms, (ii) incorporated firms, and (iii) sole proprietorships.[22] The results are shown in Figure 9a. 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, there is no statistically or economically significant increase in the number of new firms in the aggregate. The 2SLS estimates are shown in columns 1–3 of Figure 9c. A 1 pp increase in refugee to native ratio increases the number of new corporations by 1.8% and decreases the number of sole proprietorships by 0.4%. These two effects cancel each other in the aggregate, which leads to a null result on total firm formation. These results suggest that refugees increased the number of new, productive firms and decreased the number of new, less productive firms. In the aggregate, a one standard deviation increase in distance exposure, which is associated with a 5% increase in labor supply, leads to no changes in new firm formation in the Turkstat data.

To provide more evidence for this change in the productivity distribution of new firms, I utilize the GBS data and separate firms into three groups based on their participation in international trade: non-traders, exporters, and importers.[23] The intuition is that firms participating in international trade are more productive (Melitz, 2003). Hence, the existence of demand and informalization effects would imply that we should observe larger effects on trader firms and smaller, even null effects on non-trader firms. Following the same empirical strategy, I first estimate the reduced form using equation 3, 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 9b. Refugees cause significant increases in the number of both exporter and importer firms and do not change the number of non-trader firms. The 2SLS estimates are shown in columns 4–6 of Figure 9c. A 1 pp increase in the 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 does not impact the formation of non-trader firms in a statistically significant way.

Refugees' null effects on firm entry in the Turkstat data and on non-trader firm entry in the GBS data are even more surprising considering that refugees increase the local population substantially, which should create more firms via market size effects (Seim, 2006). Appendix Section C.1.2 shows that the more populous regions in Turkey have more firm creation. It further shows that refugees substantially increase the total population while not causing a significant decrease in the native population.

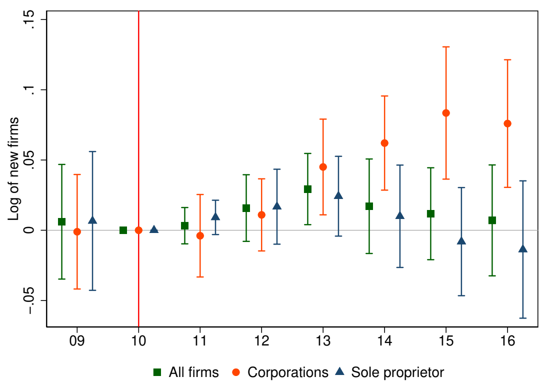
The heterogeneous effects on the number of new firms across firm types are consistent with a positive effect of immigration on firm entry and an escape to informality among less productive firms. Alternative explanations must rationalize why low-skill immigrants increase the number of productive firms,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

Without data on informal firms, I cannot credibly conclude that the informal refugee labor

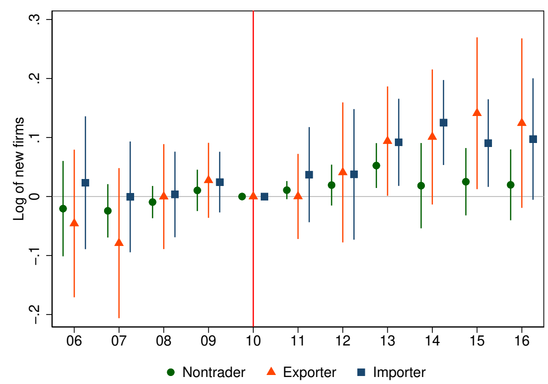
[22] Since there are only two periods before treatment, I do not adjust for linear trends.

[23] A firm is an exporter (importer) if it appears for at least once in the exports (imports) data during its lifetime.

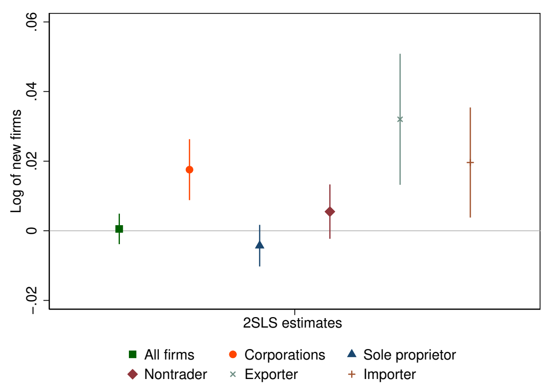
Figure 9: Refugees' effects on formal firm entry



(a) Tax status



(b) Trade participation



(c) 2SLS estimates

Notes: The points in Panels A and B represent the estimated effects of event time shown in equation 3. The 2SLS estimates in Panel C come from the IV design in equation 5. Standard errors are clustered at the nuts2 region level. The 95% confidence intervals are plotted.

supply has incentivized firms to remain unregistered. However, to make as much progress as possible without such data, I study refugees' effects on electricity consumption, which is a commonly used indicator to measure informal firm activity (La Porta and Shleifer, 2014). Figure C.1 displays the results. A 1 pp increase in refugee/native ratio increases the regional electricity consumption by 0.8%. Put differently, although refugees did not lead to significantly more firm formation in the aggregate, they caused a large increase in electricity consumption, which would be consistent with more firm activity in the informal sector.

Importantly, I do not claim that the refugee shock had no effect on new formal firm formation. I show that there are more new entrants that are incorporated or participate in international trade. However, I also demonstrate that the number of less productive firms, such as non-traders, does not increase, and the number of very small firms (sole proprietorships) appears to decrease. These results are consistent with established findings on firm entry, such as increases in the number of large firms (Altındağ et al., 2020) and firms with foreign founders (Akgündüz et al., 2018), but extend these findings to establish a theoretically meaningful pattern. The number of low-productivity firms does not increase while the number of high-productivity firms does increase. Since most new firms are typically not highly productive, I cannot reject the null hypothesis that the total number of new firms increases in the aggregate. This occurs despite significant increases in measures of total firm activity, such as electricity consumption. The accumulated evidence strongly indicates an escape to informality by small firms.

The next section introduces a model that rationalizes my main empirical findings, including the results on the extensive margin. However, due to a lack of data on informal firm entry, I abstract away from extensive margin adjustments during model estimation.

6 Theory

This section section serves two purposes. First, it rationalizes my main empirical results through a model that incorporates informal labor, firm entry, and natives' labor supply decision between salaried and non-salaried jobs. This provides a framework for thinking about the effects of modern refugee crises where host countries have sizeable informal sectors. Second, it estimates the model to quantify the role that the lack of work permits played in driving the empirical results.

Section 5 shows that the arrival of low-skill refugees who could only work informally resulted in the following changes in the labor market.

1. Low-skill natives lost both informal and formal salaried jobs.
2. High-skill natives are not affected.
3. These effects are true for both men and women.
4. Native job losses are concentrated in small firms.
5. Men transition into non-salaried positions after losing salaried jobs whereas women do not.
6. There is a missing mass of new small firms.

Results 1–4, which I denote as the intensive margin effects, can be easily rationalized by the canonical labor demand framework with a representative firm that can use both informal and formal workers in production. To make the exposition simple and obtain closed form solutions, I start with this framework. I later extend the model to incorporate an endogenous labor supply and firm entry margins to explain the fifth and sixth results, respectively.

Labor Demand

Labor is the only factor in production. The firm with productivity θ uses low-skill and high-skill labor in production. Low-skill labor is a CES aggregate of informal and formal workers, whereas high-skill labor can only be hired formally. This is consistent with Turkish data, where the probability of working informally is relatively low for natives with at least high-school degrees. Following Ulyssea (2018), I assume that, while the marginal cost of a formal worker remains constant, the firm faces an increasing and convex expected cost to hire informal workers. This assumption can be rationalized by the fact that larger firms are more likely to be caught (De Paula and Scheinkman, 2011). This convex cost structure also predicts that the probability of being informally employed decreases by firm size, which is empirically consistent with Turkish data.

The firm's objective function can be written as:

$$\max_{\ell_i,\ell_f,\ell_h} F(\ell_i,\ell_f,\ell_h) - \ell_i^{1+\gamma} w_i - w_f \ell_f - w_h \ell_h \tag{6}$$

where w_i is the informal wages of the low-skilled workers, $\gamma > 0$ captures the convex expected cost of hiring informal workers, w_f is the formal wages of the low-skilled, and w_h is the (formal) wages of the high-skilled. For notational simplicity, I omit the taxes on formal wages, which can thus be interpreted as gross wages. The production function F has a CES form,

$$F(\ell_i,\ell_f,\ell_h) = \theta(\eta_1 L^{\rho_1} + (1-\eta_1)\ell_h^{\rho_1})^{\frac{\alpha}{\rho_1}}$$
$$L = \left(\eta_2 \ell_i^{\rho_2} + (1-\eta_2)\ell_f^{\rho_2}\right)^{\frac{1}{\rho_2}}$$

where θ is the Hicks-neutral productivity term, $0 < \alpha < 1$ indicates a decreasing returns to scale (in labor) production function that is appropriate for studying short-run adjustments, L denotes low-skill labor, which itself is a CES aggregate of informal and formal low-skill workers, $\sigma_1 = \frac{1}{1-\rho_1}$ is the elasticity of substitution between low-skill and high-skill labor, $\sigma_2 = \frac{1}{1-\rho_2}$ is the elasticity of substitution between formal and informal labor, and η_1, η_2 denote the CES share parameters.

Equilibrium

Let $L_i^S(w_i)$, $L_f^S(w_f)$ denote the informal and formal labor supply curves of low-skill natives, respectively, and let $L_h^S(w_h)$ denote the labor supply of high-skill natives. Notice that the labor supply curve of low-skill natives in either sector is independent of the wage in the other sector. This simplifying assumption rules out workers' ability to search for both informal and formal jobs.[24]

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

[24] As explained in Section 2, the decrease in formal employment rates implies that the downward shift in the formal labor demand curve dominates a possible upward shift in the supply curve. 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.

equals labor demand for all labor types.[25]

$$\begin{aligned} L_i^S(w_i) &= L_i^D(w_i, w_f, w_h) \\ L_f^S(w_f) &= L_f^D(w_i, w_f, w_h) \\ L_h^S(w_h) &= L_h^D(w_i, w_f, w_h) \end{aligned} \tag{7}$$

The effects of an informal labor supply shock

In this model, the effect of a low-skill and informal labor supply shock on labor demand can be captured by the elasticities of labor demand w.r.t. informal wages. The following propositions describe these elasticities.

Proposition 1. *The elasticity of low-skill informal labor demand w.r.t. low-skill informal wages is given by*

$$\epsilon_{li,wi} = \frac{(1-\rho_1)(1-\rho_2-(\alpha-\rho_2)s_{lf}) - (\alpha-\rho_1)(1-\rho_2)s_{li}s_H}{(1-\rho_1)\left[(\alpha-\rho_2)((1-\rho_2)s_{li}+(1-\rho_2+\gamma)s_{lf}) - (1-\rho_2+\gamma)(1-\rho_2)\right] + s_H(1-\rho_2)s_{li}(\alpha-\rho_1)\gamma}$$

where $s_{li} = \frac{\eta_2 l_i^{\rho_2}}{\eta_2 l_i^{\rho_2} + (1-\eta_2) l_f^{\rho_2}}$ *is the informal labor share among low-skill workers, and* $s_{lf} = 1 - s_{li}$ *is the formal labor share among low-skill workers, and* $s_H = \frac{\eta_1 \ell_h^{\rho_1}}{\eta_1 L^{\rho_1} + (1-\eta_1)\ell_h^{\rho_1}}$ *is the share of high-skill workers in the economy. Moreover, this elasticity is always negative:* $\epsilon_{l_i,w_i} < 0$.

All derivations and proofs are in Appendix Section D. This proposition simply states that the labor demand for informal workers slopes downward. As the wages of informal labor decrease, firms demand more informal workers. Notice that when the native labor supply curve is upward sloping (i.e., not perfectly inelastic), as informal wages go down, some natives would "lose" informal jobs, while the total number of informal workers in the economy would increase.

The next proposition describes the change in the labor demand for formal, low-skill workers.

Proposition 2. *The elasticity of low-skill formal labor demand w.r.t. low-skill informal wages is given by*

$$\epsilon_{lf,wi} = \frac{(\alpha-\rho_2)(1-\rho_1) - (1-\rho_2)(\alpha-\rho_1)s_H}{(1-\rho_2-(\alpha-\rho_2)s_{lf})(1-\rho_1) - (\alpha-\rho_1)(1-\rho_2)s_{li}s_H} s_{li}\epsilon_{li,wi}$$

Moreover, if $\rho_2 > max\{\rho_1, \alpha\}$, *then this elasticity is positive.*

Proposition 2 states that, when informal and formal labor are highly substitutable, then a decrease in informal wages causes the firm to demand fewer formal workers.[26] To grasp the logic

[25] Note that informal and formal versions of low-skill natives can differ, even absent binding minimum wages, due to (i) the share parameter η_2, the increasing marginal cost of informal labor, and differences in labor supply of natives endowed with informal and formal labor.

[26] A comparable qualitative prediction is developed in Delgado-Prieto (2024), who incorporates a CRTS production function (in labor) with imperfect competition, where the price of the final good is determined by product demand, into a framework similar to Ulyssea (2018). In his model, increased informal employment can reduce existing workers' productivity by lowering the price. My model predicts a decline in formal labor demand through a different mechanism, is more parsimonious and does not require additional free parameters, which allows me to estimate the model and run counterfactuals.

underlying this result, it is easier to think about the case where $\rho_1 = \alpha$, which detaches the upper nest's effect on low-skill natives' productivity. Consider how the marginal productivity of a formal worker shifts upon the employment of an informal worker. In the case of a CRTS production function ($\alpha = 1$) and formal and informal workers not being perfect substitutes ($\rho_2 < 1$), hiring an informal worker makes formal workers *more* productive due to the Q-complementarity between workers. Consequently, the firm demands more formal labor, leading to a negative elasticity of formal labor demand, $\epsilon_{L_f,w_i} < 0$. However, as α decreases, hiring an additional worker incurs productivity losses for the rest of the workers due to decreasing returns. If α is small enough (i.e., $\alpha < \rho_2$), then the productivity loss from technological constraints (e.g., capital being constant in the short run) overpowers the productivity gain from the Q-complementarity between workers. Consequently, an informal labor supply shock that reduces informal wages can incentivize firms to substitute formal workers with informal workers.

Propositions 1 and 2 describe the changes in labor demand for low-skill workers. The next proposition describes the change for high-skill workers.

Proposition 3. *In response to an informal wage shock, demand for high-skill workers evolves according to*

$$\epsilon_{H,w_i} = \frac{(\alpha - \rho_1)s_L}{1 - \rho_1 - (\alpha - \rho_1)s_H}\epsilon_{L,w_i}$$

where $\epsilon_{L,w_i} = s_{li}\epsilon_{l_i,w_i} + s_{lf}\epsilon_{l_f,w_i}$ *is the elasticity of low-skill labor demand w.r.t. informal wages and* $s_L = 1 - s_H$ *is the share of low-skill labor in production.*

This proposition shows that the labor demand elasticity of high-skill natives ϵ_{H,w_i} cannot be signed by the model, which leaves the effect of immigrants on high-skill natives an empirical question. First, the elasticity of low-skill labor demand w.r.t. informal wages is always negative. As the costs of low-skill labor decrease, firms demand more of it. How the increase in low-skill labor impacts the demand for high-skill workers depends on the substitutability between low-skill and high-skill workers. When they are highly substitutable (i.e., $\rho_1 > \alpha$), then the elasticity is positive: a decrease in informal, low-skill wages would cause firms to replace their high-skill workers with low-skill workers. When they are complements, the opposite happens. Recall that I do not find significant changes in the labor demand for high-skill workers, which implies that $\rho_1 \approx \alpha$.

Together, Propositions 1–3 explain the first two empirical findings; $\rho_2 > \max\{\rho_1, \alpha\}$ and $\rho_1 \approx \alpha$ result in low-skill natives losing jobs in both sectors and high-skill natives being mostly unimpacted. The third empirical finding is that these effects are true for both men and women. This can be easily incorporated by introducing a third layer in the production function, where each type of worker (informal, formal, or high-skill) can be a CES aggregate of men and women. So long as men and women are highly substitutable in production among low-skill labor, the model would generate that the arrival of low-skill informal immigrants (men or women) would result in low-skill men and women losing informal and formal jobs.

The fourth empirical finding is that native job losses in the formal sector are concentrated in small firms. The next proposition explains why.

Proposition 4. *If $\rho_1 \approx \alpha$, then the elasticity of formal labor demand for low-skill workers w.r.t. informal wages is larger in absolute value for less productive firms:* $\frac{\partial |\epsilon_{l_f, w_i}|}{\partial \theta} < 0$

To see the intuition behind this result, notice that the productivity θ implicitly enters into elasticities through the labor shares: s_i, s_f, and s_h. The more productive the firm, the more it wants to produce, which it can only do by hiring more labor. Because of the increasing costs of hiring informal workers, the firm relies more on formal labor as it gets bigger: $\frac{\partial s_f(\theta)}{\partial \theta} > 0$. Therefore, larger firms rely less intensively on informal labor. When informal wages go down, the savings are relatively smaller for larger firms. Consequently, the firm's response to hiring more or less formal workers becomes weaker.

In summary, a simple model of informal labor demand using the canonical labor demand framework can explain the effects on the intensive margin. Next, I incorporate the extensions that help explain natives' escape to non-salaried work and the missing mass of new small formal firms.

6.1 Extension 1: Incorporating Firm Entry

To explain the missing mass of new small firms, I develop an equilibrium model where firms can exploit both the extensive and intensive margins of informality similar to Ulyssea (2018). Because I use the extension to make only qualitative predictions, I substantially simplify his framework to make it more tractable while relaxing his assumption that informal and formal workers are perfect substitutes. Still, firm entry prevents me from obtaining closed-form solutions. Therefore, I relax the functional form assumptions of the previous section and rely on comparative statics to show how firms react to immigration-induced wage shocks.

Firms are heterogeneous and indexed by their individual productivity θ. They produce a homogeneous good using labor as their only input. Product and labor markets are competitive, and formal and informal firms face the same prices. Similar to the baseline model, there are two skill types: low-skill workers who work either formally or informally and high-skill workers who can work only formally.

Firms

Both formal and informal firms have access to the same technology. The output of a given firm with productivity θ is given by $y(\theta, \ell) = \theta F(\ell)$, where $\ell = \{\ell_i, \ell_f, \ell_h\}$ is a vector of labor types hired by the firm. The function $F(\cdot)$ is assumed to be increasing, concave, and twice continuously differentiable in each of its inputs. For example, the nested CES form I employed in the baseline model would fit this criterion, but the results generalize to a wider range of production functions.

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

$$\pi_i(\theta, w_i) = \max_{\ell_i} \{\theta F(\ell_i, 0, 0) - w_i \tau_i(\ell)\} \tag{8}$$

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

Formal incumbents follow a similar structure to the representative firm in the baseline model.[27] They can hire workers formally but also face an increasing and convex expected cost to hire informal workers, which is summarized by the strictly convex function $\tau_{fi}(\cdot)$, $\tau'_{fi}, \tau''_{fi} > 0$. Formal firms' profit function can be written as follows:

$$\pi_f(\theta, w_i, w_f, w_h) = \max_{\ell_i, \ell_f, \ell_h} \theta F(\ell_i, \ell_f, \ell_h) - \tau_{fi}(\ell_i) w_i - w_f \ell_f - w_h \ell_h \tag{9}$$

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.

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 can differ across sectors. If firms enter either sector, they can hire labor to produce and sell the final good in period 2.

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

$$V_s(\theta, w_i, w_f, w_h) = \pi_s(\theta, w_i, w_f, w_h) \quad ; \, s \in \{i, f\}$$

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 θ decides to:

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

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

$$\begin{aligned} V_i(\overline{\theta_i}, w_i, w_f, w_h) &= E_i \\ V_f(\overline{\theta_f}, w_i, w_f, w_h) &= V_i(\overline{\theta_f}, w_i) + (E_f - E_i) \end{aligned} \tag{10}$$

[27] I abstract away from the fact that formal and informal firms also differ in their taxes. Introducing corporate taxes would not change the results in a meaningful way.

Figure 10: ZPC and free-entry

$$0 \underbrace{\qquad\qquad\qquad}_{\text{No entry}} \overline{\theta_i} \underbrace{\qquad\qquad\qquad}_{\text{Informal Entry}} \overline{\theta_f} \underbrace{\qquad\qquad\qquad}_{\text{Formal Entry}} \infty$$

where $\overline{\theta_i}$ and $\overline{\theta_f}$ are the productivity of firms that are at the margin of entering into informal and formal sectors, respectively. For future reference, I refer to these firms as the marginally informal and marginally formal firms, respectively. The least productive entrepreneurs with productivity $\theta < \overline{\theta_i}$ choose not to enter. Firms with productivity $\theta \in [\overline{\theta_i}, \overline{\theta_f}]$ are productive enough to make positive profits and prefer the informal sector. The more productive firms with productivity $\theta > \overline{\theta_f}$ 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 10. The mass of new formal firms is given by $(1 - \overline{\theta_f})\mathcal{M}$.

Equilibrium

Let $L_m^S(w_m)$, $m \in \{i, f, h\}$ denote labor supply curves for the low-skill informal, low-skill formal, and high-skill formal workers. As in the baseline model, I assume that natives are either low-skill or high-skill, and low-skill natives are endowed with either informal or formal labor. In equilibrium, labor markets must clear: wages are such that labor supply equals labor demand in all sectors.

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 from the Walras' Law.

Effects of an Informal Labor Supply Shock on Firm Entry

In this model, immigrants can impact firm entry in two ways. First, they alter the informal labor supply, which causes an exogenous change in wages. Second, they can alter the pool of potential entrants. This can take place via changes in (2a) the mass of potential entrants $\mathcal{M}$, and (2b) the cdf of the productivity distribution G. These can be justified, for example, by immigrants' entrepreneurial activities or by natives' substitution from salaried work to becoming employers. For comparative statics, I focus on the first two channels: a decrease in informal wages and an increase in the mass of potential entrepreneurs. The following proposition formalizes how the number of firms entering the informal and formal sectors changes through these two mechanisms.

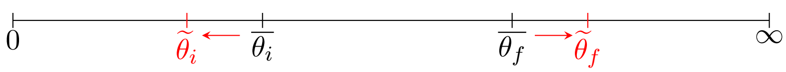
Proposition 5. *Consider the marginally formal firm $\overline{\theta_f}$. Let $\ell_i^D(\overline{\theta_f})$ and $\ell_{fi}^D(\overline{\theta_f})$ denote the number of informal workers hired by the informal and formal versions. If the informal firm hires more informal workers at the baseline equilibrium, then a decrease in informal wages incentivizes the marginal firm to remain in the informal sector. Formally, $\ell_i^D(\overline{\theta_f}) > \ell_{fi}^D(\overline{\theta_f}) \Rightarrow \frac{\partial \overline{\theta_f}}{\partial w_i} < 0$. More-*

over, the decrease in informal wages necessarily lowers the productivity threshold of the marginally informal firm: $\frac{\partial \overline{\theta_i}}{\partial w_i} > 0$.

Proposition 5 is intuitive. The marginally formal firm $\overline{\theta_f}$ has to be indifferent between entering the informal and formal sectors. If the informal version of this firm has more informal workers, then the envelope theorem dictates that a decrease in informal wages benefits the informal version more than the formal version, resulting in the firm entering into the informal sector. Let $\widetilde{\theta_f}$ denote the productivity of the new marginally formal firm. As a result of the informal wage decrease, $(\widetilde{\theta_f} - \overline{\theta_f})\mathcal{M}$ many firms disappear from the formal sector.

The second part of this proposition states that the productivity requirement to enter the informal sector $\overline{\theta_i}$ decreases as the informal wage decreases. This is intuitive since lower wages increase profits. The marginal firm that did not make any profit after paying the entry cost starts making profits as an informal firm. Naturally, the new marginal firm $\widetilde{\theta_i}$ is less productive than the original marginal firm: $\widetilde{\theta_i} < \overline{\theta_i}$. Figure 11 depicts the change in productivity thresholds in response to a decrease in informal wages.

Figure 11: Effect of decrease in informal wages on the extensive margin



Under what conditions would the informal version of the marginally formal firm hire fewer informal workers than its informal version? To answer this question, assume that hiring informal workers created the same costs for informal and formal firms: $\tau_i(x) = \tau_{fi}(x)\ \forall x \in \mathcal{R}_+$. In this model, any complementarity between formal and informal workers would induce the formal firm to employ more informal workers. In contrast, differences in the hiring costs of informal workers, $\tau_i(x) < \tau_{fi}(x)\ \forall x \in \mathcal{R}_+$, would cause the formal firm to lower its demand for informal workers. These two forces oppose each other. When the difference between hiring informal workers is stronger than the complementarity between informal and formal workers, an informal labor supply shock can induce the marginal firms to remain in the informal sector.

Absent changes in the pool of potential entrants, the change in informal wages would increase informal firm entry and decrease formal firm entry. However, I document no change in formal firm entry and a missing mass of new small firms. This can be rationalized by a counteracting change in the mass of new potential firms. An increase in the mass of new potential entrants, all else equal, increases firm entry in both sectors. Consequently, while the decrease in informal wages lowers formal firm entry by making less productive firms prefer the informal sector, the increase in the mass of potential entrants increases formal firm entry. The null effect on new formal firm formation implies that these two forces nullify each other. The testable implications are a missing mass of new formal firms and an increase in the number of new productive firms, both of which I document in the data.

Note that Propositions 1–4 still apply to formal firms in this model. The difference is that the aggregate effects on labor demand depend on an integral over the firms choosing to enter into the informal and formal sectors, which are endogenously determined. Consequently, no closed-form solution can be provided for the changes in the aggregate labor demand.

6.2 Extension 2: Endogenous Labor Supply Decision Between Salaried and Non-salaried Jobs

Men's escape to non-salaried work upon losing salaried jobs means that men's outside option to salaried employment is non-salaried employment rather than unemployment. This is a straightforward intuition that can be easily formalized by the home production framework applied to non-salaried jobs (Gronau, 1977). Individuals of each skill type are endowed with time T, which they can use to allocate between leisure l, salaried employment h_s which pays constant wages w_s, and non-salaried employment h_n. Production from non-salaried work is given by the concave function g. Home production and market goods are perfect substitutes. Consumers get utility $U(c, l)$ from leisure and consumption . They consume what they produce at home or buy at the market $c = g(h_n) + w_s h_s$, and are subject to a time constraint: $T = l + h_n + h_s$.

Assuming an interior solution, which can be guaranteed with functional form assumptions, we get $g'(h_n) = w_s$: people work in non-salaried jobs until the marginal return from non-salaried work equals salaried work. As wages fall, for example, due to an immigration shock, people transition to non-salaried jobs as g is strictly concave. How much people transition to non-salaried jobs depends on the (inverse of) curvature of the non-salaried production function g. The fact that in Turkey, low-skill men transition to non-salaried jobs and women do not can then be explained by the differences in the curvatures of the g function between men and women.[28]

6.3 How Substitutable are Formal and Informal Labor in Production?

I motivated the empirical analysis by arguing that informal refugees' impact on the formal labor markets is ambiguous. Empirical results indicate that Syrian refugees caused natives to lose both informal and formal jobs, which implies that these two types of labor are largely substitutable in production. It is of general interest to quantify their substitutability, which is governed by ρ_2. Here, I briefly describe how the IV estimates and certain moments from the data help identify the model parameters. The details of the model estimation are shown in Appendix Section E.

In this model, the sole means by which firms can augment their output is by increasing their workforce, since labor constitutes the exclusive input in the production process. Consequently, the distinction between larger and smaller firms hinges entirely upon disparities in their productivity. More productive firms choose to expand their workforce. The parameter γ, which governs the marginal cost of employing informal workers, predominantly hinges on the extent to which larger firms opt for formalization at the intensive margin. For all types of firms, the share parameter η_2 is linked to the relative productivity of formal and informal workers and, thus, is determined by the proportion of informal workers in the overall economy. The elasticity of substitution between

[28]Unfortunately, Turkish labor force surveys do not collect earnings information from non-salaried workers. Therefore, I cannot test this assumption.

informal and formal workers is primarily dictated by demand elasticities for low-skill natives. The null effect on high-skill natives pins down the substitutability between low-skill and high-skill labor: $\rho_1 \approx \alpha$.

Using the IV estimates and moments from the data, I estimate ρ_2 to be around 0.9, which implies an elasticity estimate, $\sigma_2 = \frac{1}{1-\rho_2}$, of around 10. To the best of my knowledge, this is one of the first papers to estimate this elasticity. This relatively high elasticity is consistent with the Turkish context, where informal employment is often in the same sectors and even in the same firms as formal employment. It also supports the assumption of perfect substitutability between informal and formal workers in the recent structural literature on the informal sector (Ulyssea, 2018, 2020).

6.4 The Effects of Granting Refugees Work Permits

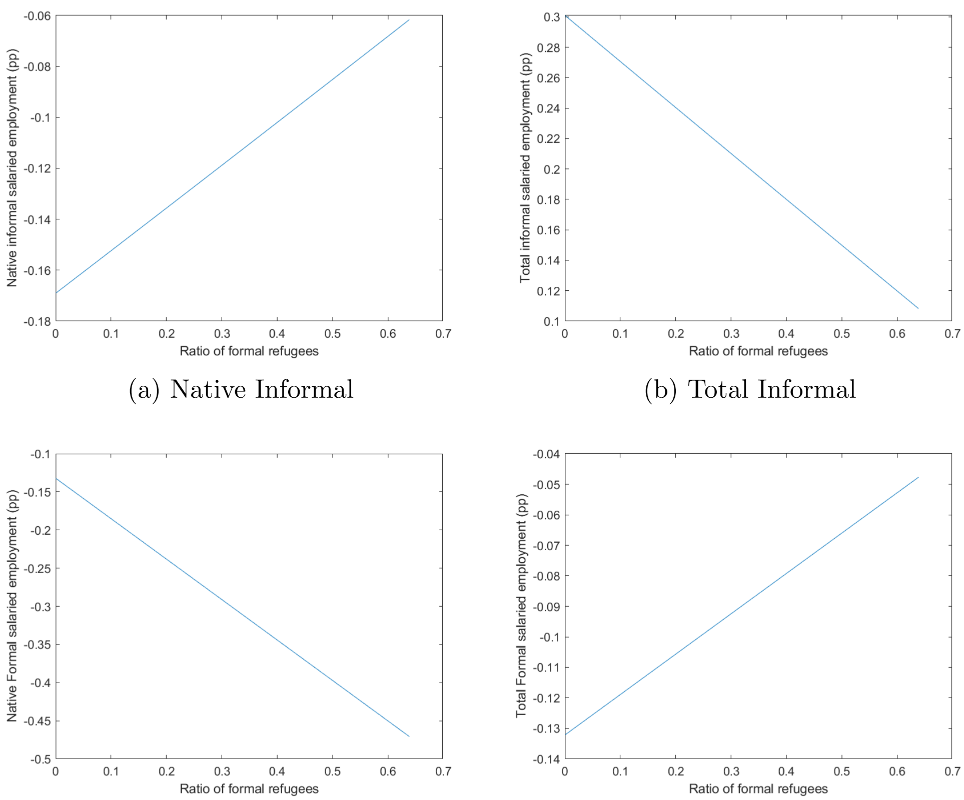
The presence or absence of work permits constitutes a pivotal distinction between immigration episodes and contemporary refugee crises. Unlike immigrants, most refugees lack formal authorization to participate in the labor markets (Clemens et al., 2018). For example, most Syrian refugees in Turkey remain without work permits as of 2024. However, this approach is not uniformly applied across nations. Colombia, for instance, adopted a phased approach by granting work permits to Venezuelan refugees in waves (Bahar et al., 2021). Furthermore, nearly all European countries extended the right to work for Ukrainian refugees (European Commission, 2022). Most recently, the United States announced its intention to provide work permits to Venezuelan refugees already residing within its borders (Hesson, 2023). Given the diverse strategies different countries employ regarding work permits and the far-reaching implications of these policies spanning multiple nations, it is imperative to comprehend the repercussions associated with providing refugees with work permits. This section studies the counterfactual outcomes if Turkey were to grant all Syrian refugees work permits. Does providing refugees with work permits hurt native workers? Does it change firms' incentives to employ informal labor?

In the baseline model, where labor is the only factor of production and labor supplies are taken as given, introducing work permits for refugees has a singular effect: it reallocates a portion of the informal labor force into the formal sector. This reallocation causes a *reduction* in the total informal labor supply in the economy, leading to an increase in informal wages and an increase in informal employment of natives. In the formal sector, assuming a binding minimum wage for simplicity, the shift in the formal labor supply curve does not affect wages. Consequently, formal employment depends exclusively on the demand for formal labor. Therefore, if there is no shift in the formal labor demand curve, the employment of refugees in the formal sector would lead to an equivalent reduction in the employment of native workers in that sector. However, since the informal wage elasticity of formal labor demand is positive (i.e., $\alpha < \rho_2$ in the model), the increase in informal wages pushes the formal labor demand curve outwards, increasing total formal employment in the economy. The magnitude of these changes depends on two factors: (1) the model parameters, estimates of which are reported in Table E.1 in the Appendix, and (2) the percentage of working refugees who can transition to the formal sector if given permits.

Let $c \in [0, 1]$ denote the ratio of refugees that are endowed with only formal labor; $c = 0$ implies that all working refugees are constrained to informal labor even with work permits. Conversely, $c = 1$ implies that all working refugees would secure formal employment if granted work permits. Unfortunately, there is no good data-driven way to estimate c. In Turkey, there are very few and highly selected refugees with work permits. Therefore, I cannot credibly estimate c from the data. Instead, I assume that refugees are weakly less formal than natives: $c \in [0, 0.64]$, which is a conservative assumption.

Figure 12 shows the counterfactual effects of a 1 pp increase in refugee/native ratio for all potential values of c. As a benchmark, if refugees had the same formality rate as natives, a 1 pp increase in refugee/native ratio would have caused a 0.061 pp decrease in native informal employment, 0.11 pp increase in total informal employment, 0.47 pp decrease in native formal employment, and only a 0.047 pp decrease in total formal employment (as opposed to the 0.13 pp decrease estimated in the empirical section). Intuitively, as more refugees can find formal jobs, fewer natives lose informal jobs, and more natives lose formal jobs.[29]

Figure 12: Effects of a 1 pp increase in refugee/native ratio with different levels of refugee informality



(a) Native Informal (b) Total Informal

(c) Native Formal (d) Total Formal

[29]In follow-up work studying the Venezuelan refugee crisis in Colombia, Bahar et al. (2024) show reduced-form evidence for this prediction.

A direct interpretation of these findings is that not providing work permits to refugees costs 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. 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 in Figure 12d and the benchmark case of refugees having the same informality rate as low-skill Turkish natives, I conclude that not providing work permits to refugees caused approximately 120 thousand formal jobs to disappear in 2016. At the time, the formal monthly minimum wage was around \$549 before tax and \$433 after tax. Assuming that all the jobs lost were paying the minimum wage, not providing work permits to refugees cost 167 million USD in personal income tax revenue to Turkey in 2016.

7 Conclusion

This paper provides an empirical and theoretical analysis of how firms and native workers respond to an informal labor supply shock, using the Syrian refugee crisis in Turkey as a quasi-experiment. The findings illuminate our understanding of the informal economy and have important policy implications.

I show that an increase in the informal labor supply due to the influx of Syrian refugees significantly impacts both the informal and formal sectors. Native salaried employment decreases in both the informal and formal sectors. Native disemployment in the informal sector can be explained by a downward-sloping labor demand curve in the informal sector. However, the native disemployment in the formal sector, despite refugees' inability to work formally, highlights that firms substitute formal workers for informal workers. Robustness checks confirm that the disemployment effects result from refugees' informal labor. Evidence suggests that this displacement happens both on the intensive margin, by formal incumbents lowering their demand for formal workers, and on the extensive margin, by new firms choosing to remain unregistered. Moreover, low-skill men who lose their salaried jobs transition into non-salaried jobs. This adjustment is economically and empirically significant, underscoring the importance of distinguishing between salaried and non-salaried employment when studying immigrants' effect on the labor market.

I rationalize the empirical findings by proposing a model of labor demand where firms can use both informal and formal labor in production. Taking the model to the data, I estimate an elasticity of substitution between formal and informal labor of around 10 and quantify the role that the lack of work permits play in driving the empirical results. I show that permits boost native employment in the informal sector while reducing it in the formal sector. However, the increase in informal wages encourages firms to hire more formal workers, ultimately creating more formal jobs. The magnitude of these changes depends on the formality rate of refugees, with significant potential benefits in terms of job creation and government tax revenue.

My results are particularly relevant for policy today, as the global refugee population continues to grow, forcing governments to decide how to manage these humanitarian crises. Despite various ethical and economic arguments for granting refugees work permits, most refugees still lack permits today. I provide a new argument based on empirical evidence and theoretical analysis: denying work permits to refugees incentivizes firms to become more informal. This generates several undesirable economic consequences, including lost tax revenue as I quantify in this paper. Additional losses include efficiency losses from incentivizing firms to remain small to avoid detection, unfair competition between small (informal) and large (formal) firms, and skill mismatches when high-skill immigrants work in low-skill industries. Future research can investigate these channels further.

This research provides valuable insights into the complex dynamics of the informal economy, the labor market effects of refugee inflows, and the potential policy implications of granting work permits to refugees. The findings challenge conventional assumptions and offer a nuanced understanding of the interactions between formal and informal sectors in the context of an informal labor supply shock.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 2010, *105* (490), 493–505.

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.

_ **, Marcel van den Berg, and Wolter Hassink**, "The impact of the Syrian refugee crisis on firm entry and performance in Turkey," *The World Bank Economic Review*, 2018, *32* (1), 19–40.

_ **, Yusuf Kenan Bağır, Seyit Mümin Cılasun, and Murat Güray Kırdar**, "Consequences of a massive refugee influx on firm performance and market structure," *Journal of Development Economics*, 2023, *162*, 103081.

Aksu, Ege, Refik Erzan, and Murat Güray Kırdar, "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 one unatives," *The Economic Journal*, 2003, *113* (488), F302–F331.

Aslan, Melih, "UNHCR to seek new help for refugees in Turkey, push resettlement," https://www.reuters.com/article/world/unhcr-to-seek-new-help-for-refugees-in-turkey-push-resettlement-idUSKCN0UU0MF/#:~:text=UNHCR%20to%20seek%20new%20help%20for%20refugees%20in%20Turkey%2C%20push%20resettlement,-By%20Melih%20Aslan&text=ISTANBUL%20(Reuters)%20%2D%20The%20United,to%20enter%20its%20sixth%20year. 2016. Accessed: 2025-06-29.

Azoulay, Pierre, Benjamin F Jones, J Daniel Kim, and Javier Miranda, "Immigration and entrepreneurship in the United States," *American Economic Review: Insights*, 2022, *4* (1), 71–88.

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.

_ , Isabel Di Tella, and Ahmet Gulek, "Formal Effects of Informal Labor and Work Permits Evidence from Venezuelan Refugees in Colombia," 2024. Available at https://github.com/ahmetgulek/ahmetgulek.github.io/blob/main/Bahar_DiTella_Gulek_Colombia.pdf.

Borjas, George J 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.

Chodavadia, Saheel A, Sari Pekkala Kerr, William R Kerr, and Louis J Maiden, "Immigrant entrepreneurship: New estimates and a research agenda," 2024. NBER WP:32400.

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, "Immigration, wages, and employment under informal labor markets," *Journal of Population Economics*, 2024, *37* (2), 55.

Demirci, Murat and Murat Güray Kırdar, "The labor market integration of Syrian refugees in Turkey," *World Development*, 2023, *162*, 106138.

Dimmock, Stephen G, Jiekun Huang, and Scott J Weisbenner, "Give me your tired, your poor, your high-skilled labor: H-1b lottery outcomes and entrepreneurial success," *Management Science*, 2022, *68* (9), 6950–6970.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J Notowidigdo, "The economic consequences of hospital admissions," *American Economic Review*, 2018, *108* (2), 308–352.

Dustmann, Christian, Tommaso Frattini, and Ian P Preston, "The effect of immigration along the distribution of wages," *Review of Economic Studies*, 2013, *80* (1), 145–173.

_ **, Uta Schönberg, and Jan Stuhler**, "The impact of immigration: Why do studies reach such different results?," *Journal of Economic Perspectives*, 2016, *30* (4), 31–56.

Elias, Ferran, Joan Monras, and Javier Vazquez-Grenno, "Understanding the effects of granting work permits to undocumented immigrants," 2025. Forthcoming at the Journal of Labor Economics.

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.

European Commission, "EU Solidarity with Ukraine," https://eu-solidarity-ukraine.ec.europa.eu/information-people-fleeing-war-ukraine/fleeing-ukraine-your-rights-eu_en 2022. Accessed: 2024-01-18.

Friedberg, Rachel M, "The impact of mass migration on the Israeli labor market," *The Quarterly Journal of Economics*, 2001, *116* (4), 1373–1408.

Gronau, Reuben, "Leisure, home production, and work–the theory of the allocation of time revisited," *Journal of Political Economy*, 1977, *85* (6), 1099–1123.

Gulek, Ahmet and Jaume Vives i Bastida, "Synthetic IV Estimation in Panels," 2025. Available at https://jvivesb.github.io/jmp/draft_siv.pdf.

_ **and Tishara Garg**, "The Effects of Immigration on Regions that Don't Receive Immigrants," 2025. Available at https://ahmetgulek.github.io/Gulek_jmp.pdf.

Hesson, Ted, "US offers work permits to half million Venezuelans already in country," 2023. https://www.reuters.com/world/us-offers-work-permits-half-million-venezuelans-already-country-2023-09-21/#:~:text=WASHINGTON%2C%20Sept%2020%20(Reuters),newly%20arrived%20migrants%20work%20legally. [Accessed: 2023-09-21].

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 2022. Accessed: 2022-06-21.

Lofstrom, Magnus and Chunbei Wang, "Immigrant entrepreneurs," *Handbook of labor, human resources and population economics*, 2022, pp. 1–32.

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.

Monras, Joan, "Immigration and wage dynamics: Evidence from the mexican peso crisis," *Journal of Political Economy*, 2020, *128* (8), 3017–3089.

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, Semsa, "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 Yasenov, "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," 2014. Available at https://sites.bu.edu/neudc/files/2014/10/paper_189.pdf.

Seim, Katja, "An empirical model of firm entry with endogenous product-type choices," *The RAND Journal of Economics*, 2006, *37* (3), 619–640.

Sevinc, Orhun, Emre Cakir et al., "Rising Labor Share in Turkey," Technical Report, Research and Monetary Policy Department, Central Bank of the Republic of Turkey 2021.

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].

Wolfers, Justin, "Did unilateral divorce laws raise divorce rates? A reconciliation and new results," *American Economic Review*, 2006, *96* (5), 1802–1820.

Xu, Yiqing, "Generalized synthetic control method: Causal inference with interactive fixed effects models," *Political Analysis*, 2017, *25* (1), 57–76.

Appendix

A Data

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, 26 NUTS-2, and 81 Nuts-3 regions. All of the analyses in the main text are conducted at the 26 NUTS-2 level to maintain consistency across different datasets unless specified otherwise.

The text combines employment into two categories: salaried and non-salaried jobs. In general, salaried jobs are more desirable than non-salaried jobs. Not surprisingly, the probability of a job being a salaried job increases with education, formality, and regional GDP.

Table A.1: HLFS Summary Statistics

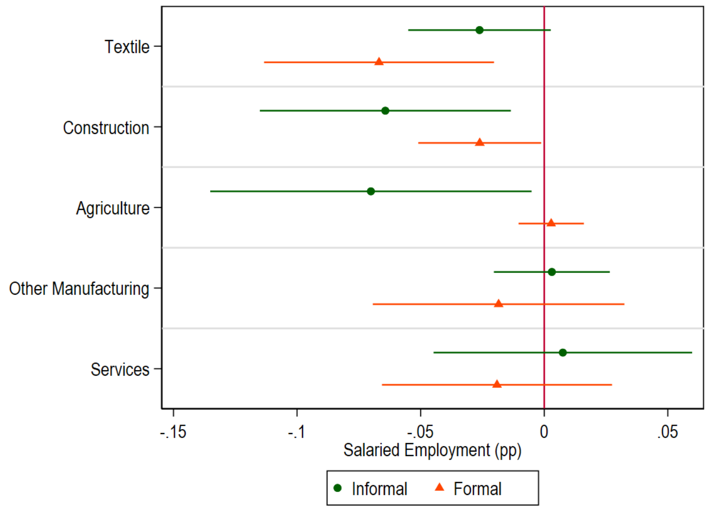
	Salaried Employment				Non-salaried Employment			
Formality	All	Informal	Formal		All	Informal	Formal	
Skill			Low	High			Low	High
Panel A: Aggregate								
Pooled	0.323	0.071	0.157	0.459	0.188	0.124	0.061	0.071
Men	0.491	0.106	0.292	0.544	0.251	0.134	0.122	0.107
Women	0.160	0.037	0.045	0.340	0.127	0.115	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. Salaried employment is defined as regular, salaried work. Non-salaried employment consists 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. Details can be found following this link: https://ilostat.ilo.org/resources/concepts-and-definitions/classification-economic-activities/

B Supporting Evidence Summarized in Section 5.1

Native employment is categorized into five groups: textile, construction, agriculture, other manufacturing, and services, following ISIC definitions. Figure B.1 shows the estimated refugee effects on low-skill natives on each category. The disemployment effects in the informal and formal sectors come mostly from refugee hiring industries. Most notably, the textile industry observes the largest decrease in formal employment. A 1 pp increase in refugee/native ratio decreases natives' informal salaried employment by 0.03 pp and formal salaried employment by 0.07 pp. Put differently, the industry that hires the most refugees lets go the most natives. Other manufacturing industries do not observe similar decreases in salaried employment. Moreover, natives lose informal and formal jobs also in construction, the second most intensely treated industry. Lastly, there is no change in formal salaried employment in services.

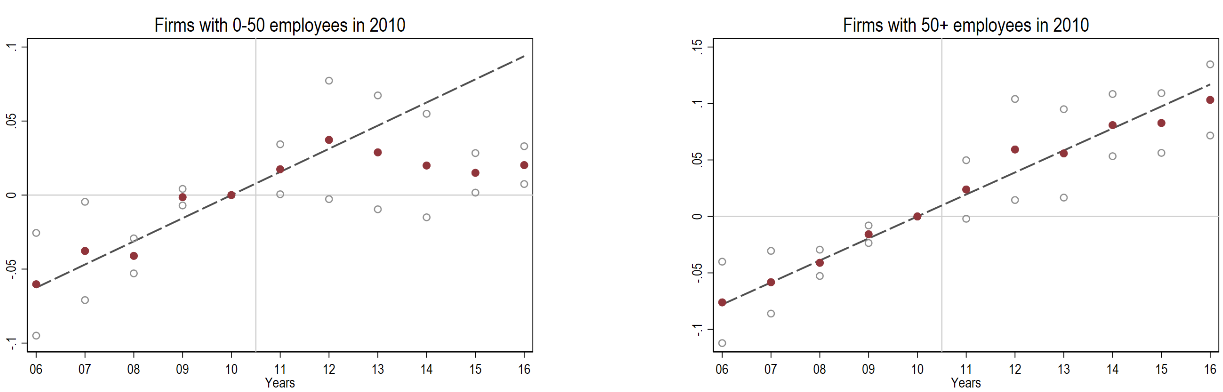
Figure B.1: Industry Heterogeneity



Notes: The 2SLS estimates come from estimating equation 5 using natives' informal, low-skill formal, and high-skill formal salaried employment rates. The first row shows the estimates using the pooled data. The second and third rows condition on men and women separately. The 95% confidence intervals are plotted.

Figure B.2 shows that the decrease in formal employment in Textile comes solely from small firms and not from large firms. Interestingly, I document similar pre-trends for both small and large firms. Both types of firms grew by around 6-8% more in regions closer to the border compared to regions away from the border between 2006–2010. However, small firms start deviating from their trend starting in 2013, whereas large firms continue theirs.

Figure B.2: Refugees' effect on Textile Firms



(a) 0-50 employees

(b) 50+ employees

Online Appendix for "Formal Effects of Informal Labor" by Ahmet Gulek

To avoid confusion with the Appendix included in the manuscript, the Online Appendix starts from Section C.

C Additional Supporting Empirical Evidence

C.1 Supporting Evidence Summarized in Section 5.4

C.1.1 Electricity Consumption

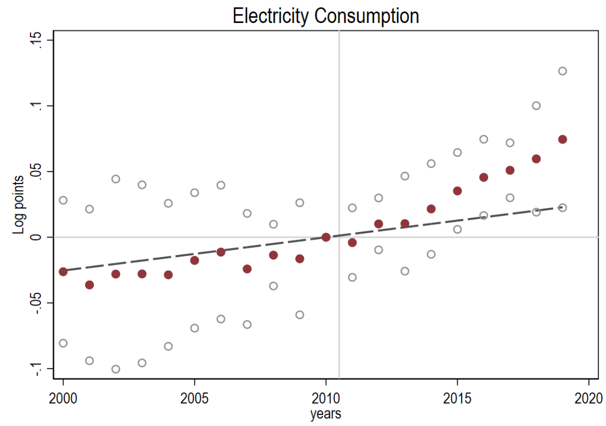
Section 5.4 of the main text investigates whether informal immigration impacts firms' decisions to formalize on the extensive margin; i.e., register with the tax authorities. It documents a change in the productivity distribution of new formal firms: a decrease in the number of less productive firms and an increase in more productive firms. It argues that the missing mass of new small formal firms is indicative of less productive entrepreneurs choosing to remain unregistered to have easier access to informal labor. If true, this would be an additional effect of an informal labor supply shock. However, the lack of credible data sources on unregistered firms in Turkey prevents testing whether the number of informal firms has increased.

Without data on informal firms, I cannot credibly conclude that the informal refugee labor supply has incentivized firms to remain unregistered. 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). Data on electricity consumption at the province level comes from Turkstat. For consistency with the rest of the paper, I perform the analysis at the NUTS2 level. I estimate the nonparametric and parametric event study designs shown in equations 3 and 4. Figure C.1 shows the point estimates from the nonparametric design, and the linear trend from the parametric design. The distance exposure is associated with significant and positive deviations from the trend after 2015. A one standard deviation increase in the instrument is associated with a 3.8% increase in electricity consumption in 2016. Put differently, 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.

C.1.2 Refugees' effect on Native population

In the main text, I argue that refugees' null effect on the creation of non-trader firms is highly suggestive of new firms choosing to remain informal. This is because there is a well known relationship between firm entry and market size, and therefore an increase in population should cause more firm creation. An alternative hypothesis could be that refugees decrease the native population of the host regions, e.g., by increasing out-migration or decreasing in-migration. If this effect was

Figure C.1: Event study design on electricity consumption



Notes: The points in each figure represent the estimated effects of event time shown in equation 3. The hollow circles present the 95 percent confidence intervals. The dashed line represents the estimated pre-2010 linear relationship between outcome and instrument * event time from the parametric event study in equation 4 with the level normalized to match the nonparametric estimates.

large enough, refugees could decrease new firm creation simply by reducing native population. In this section, I show evidence againts this alternative hypothesis. First, I document the relationship between native population and firm entry. Second, I show that refugees decrease in-migration and increase out-migration of natives my economically insignificant amounts. Consequently, refugees do not impact the native population in the time period of study.

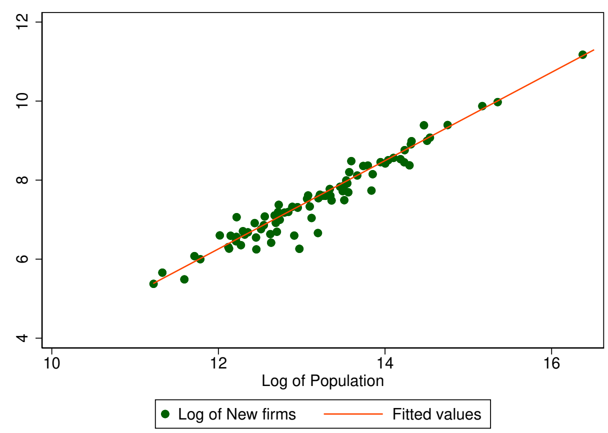
C.1.3 Relationship between population and firm entry

In Turkey, population and number of new firms is strongly correlated. Figure C.2 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. Across provinces, a 1% increase in native population is associated with a 1.1% increase in new firm entry per year. This suggestive correlation 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 the natural logarithm of number of new firms on the natural logarithm of local population while controlling for province and year fixed effects in the pre-period result in an elasticity estimate of around 0.75, which is still large.

C.1.4 Refugees' Null effect on native population

In this subsection, I show that refugees have only a minor effect on in-migration and out-migration of natives. Consequently, they lead to no significant change in the native population. If anything, the treated regions keep observing a growth in their native population due to higher

Figure C.2: Market size and firm entry in 2009

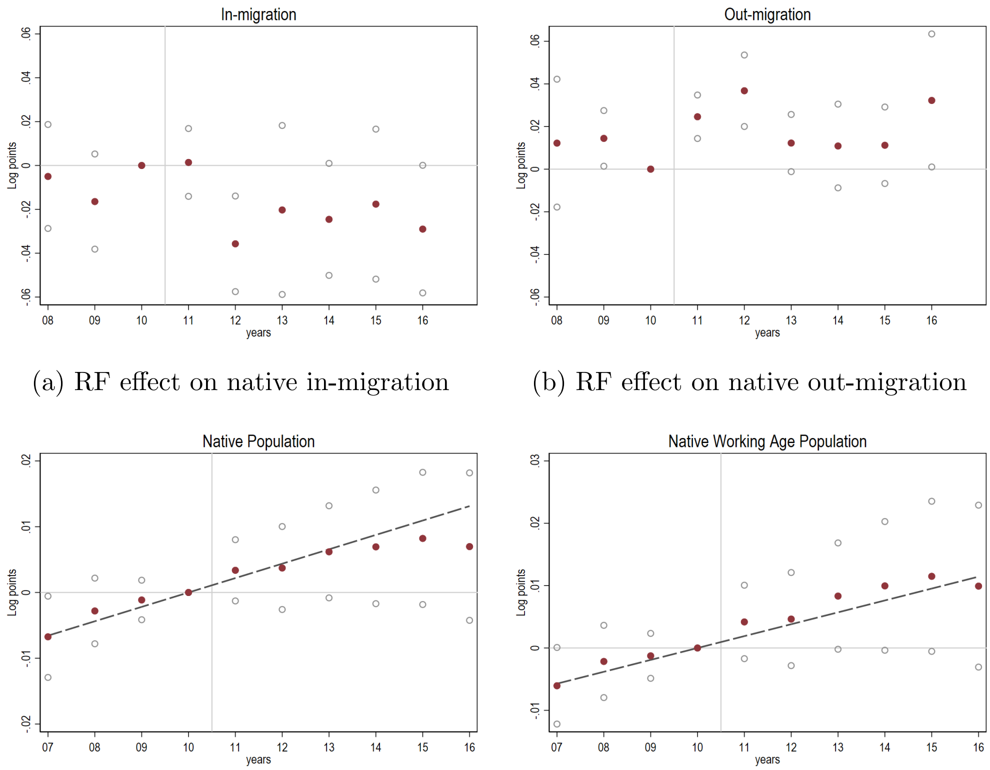


growth rates. To show this, I estimate the nonparametric event study design shown in equation 3 of the main text where the outcome variable is the natural logarithm of the amount of in-migration and out-migration at the region-year level. Panels A and B of Figure C.3 shows the results. we see that the provinces closer to the border observed statistically significant changes in both in-migration and out-migration. The effects are 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. Panels C and D of Figure C.3 plot the same event study figures on the natural logarithm of the native population and working age native population, respectively. We see that the regions closer to the border were observing larger increases in their populations in percentage terms even before the refugee crisis began. However, the crisis did not alter this pre-existing trajectory. The parametric linear trend falls within the nonparametric estimates in all years.

Figure C.3: Nonparametric Event study figures on native population



(a) RF effect on native in-migration

(b) RF effect on native out-migration

(c) RF effect on native population

(d) RF effect on working age population

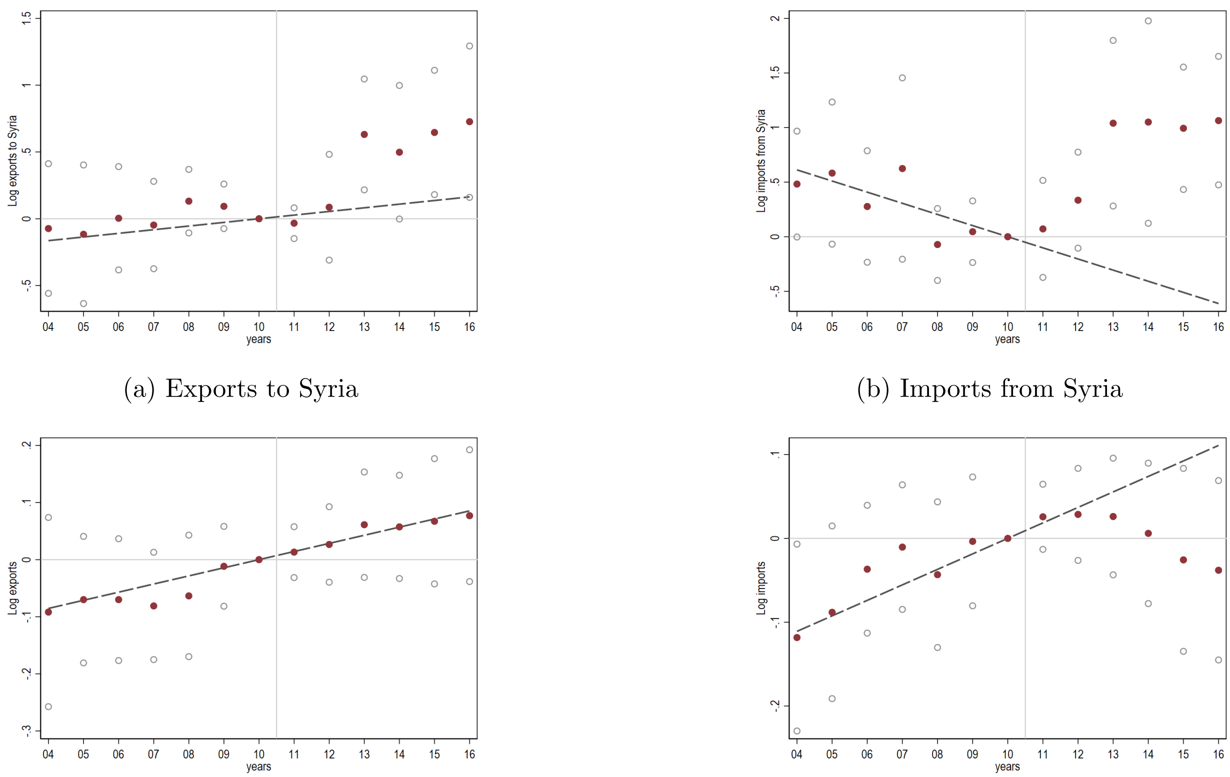
C.2 Other Supporting Evidence

Trade-related confounders

I rely on a spatial IV-DiD strategy to identify the effects of Syrian refugees on labor markets. I use a distance-based instrument, which boils down to comparing regions close to the border with regions that are further away. This empirical strategy assumes that the Syrian war's impact on the Turkish local labor markets, if any, should be orthogonal to the distance from the border. This could fail if Syria were a major trade partner of border regions and the war had significantly disrupted the trade flows. To investigate this, I calculated the trade flows between Turkish regions and Syria and the rest of the world from Turkstat's customs data. In particular, for each region-year cell, I calculated the total amount of exports to Syria, total exports to other countries, total imports from Syria, and total imports from other countries. I then estimate the nonparametric and parametric event study designs shown in equations 3 and 4, where the outcome variables are the natural logarithm of trade flows.

Panels A, B, C, and D of Figure C.4 plot the results. Panels A and B show that regions close to the border do not observe significant decreases in imports from and exports to Syria. If anything, exports to and imports from Syria actually increase after 2011. This evidence rules out a negative trade shock causing native disemployment in the border regions. Moreover, the trade relations with Syria were not significant enough to disrupt the labor markets. This can be seen in Panels C and D, which show the effect of distance on total exports and imports. Despite regions closer to the border observing increases in trade with Syria, total exports remain unaffected, and total imports decrease by a small amount. The latter is likely a causal effect of the refugee labor supply, which lowers the production costs of local goods. Overall, the evidence strongly suggests that the Syrian Civil War did not cause a significant trade shock to Turkey that can explain my findings.

Figure C.4: Event study estimates on exports and imports



(a) Exports to Syria

(b) Imports from Syria

(c) Total Exports

(d) Total Imports

Notes: The points in each figure represent the estimated effects of event time shown in equation 3. The hollow circles present the 95 percent confidence intervals. The dashed line represents the estimated pre-2010 linear relationship between outcome and instrument * event time from the parametric event study in equation 4 with the level normalized to match the nonparametric estimates.

C.3 Synthetic IV Adjustment

This section describes the robustness checks of the main results using the Synthetic IV (SIV) methodology. SIV is a non-parametric method that combines the instrumental variable strategy with synthetic controls. I provide a brief description of how the method works here, and refer the reader to Gulek and Vives-i Bastida (2025) for a full treatment.

In summary, the procedure is as follows. Let {Y,R,Z} denote the dataset at hand, where Y is the outcome, R is the treatment, and Z is the instrument. First, find synthetic control (SC) weights for each unit, regardless of its treatment status, by solving the standard synthetic control program. Then, use these weights to generate synthetic data, which includes the outcome $\hat{Y}_{it}^{SC}$, treatment $\hat{R}_{it}^{SC}$, and instrument $\hat{Z}_{it}^{SC}$. Then, subtract the synthetic data from the real data to obtain *debiased* data ($\tilde{Y}_{it} = Y_{it} - \hat{Y}_{it}^{SC}$, $\tilde{R}_{it} = R_{it} - \hat{R}_{it}^{SC}$, $\tilde{Z}_{it} = Z_{it} - \hat{Z}_{it}^{SC}$). Finally, estimate the desired model using the *debiased* data in the post period. For example, a simple implementation would be to use the pre-treatment values of the outcome Y to calculate the SC weights, and then estimate the IV model using the debiased data in the post period by employing 2SLS.

Intuitively, matching on pre-trends addresses the pre-trend problem. However, it does not address the fact that immigrants can choose their location based on contemporaneous economic shocks. This is still addressed by the instrument Z. Put differently, SIV addresses the unobserved confounding problem via synthetic control and the endogeneity problem via the instrument.

To implement the algorithm, I first demean the outcomes of interest by subtracting their pre-treatment mean values before solving the synthetic control program. This is equivalent to adding a constant term in the SC problem. Then, I solve the standard synthetic control problem by matching on the demeaned pre-treatment values between 2004–2010. To provide robustness checks for this robustness check, I also do a backtesting exercise and show results where the training was done using values between 2004–2007. Dividing the pre-treatment period into a training and a testing set enables me to visually check that I do not overfit. After calculating the debiased data using these weights, I estimate refugees' effect using the debiased data in the post period 2011–2016.

Figure C.5a shows the reduced form effect on the salaried employment rates of low-skill natives.[30] Notice that despite matching on the data between 2004–2007, SIV corrects for the pre-trend between 2008–2010. This implies that the algorithm captures the signal in the data. Using either SIV estimate finds significant declines in salaried employment rates in the post period.

Figure C.5b compares the IV estimate without adjusting for linear trends, the IV estimate adjusting for the linear trend, and the Synthetic IV estimates on the salaried employment rate and formal salaried employment rate of low-skill natives. Both the IV estimate with linear trend and the SIV estimates find that informal immigrants displace low-skill natives in the formal sector.

D Model Derivations and Proofs

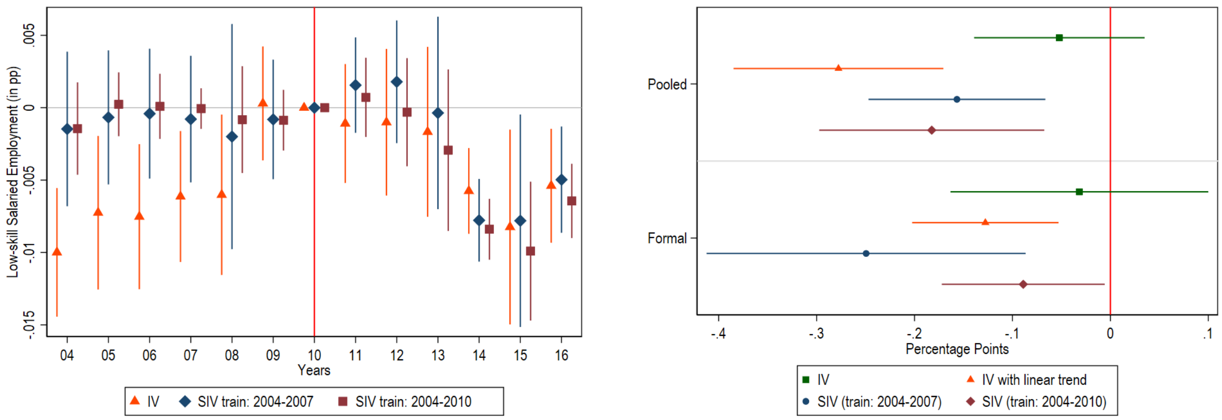
D.1 Model of informality with only one skill type

An earlier version of this paper employed a model where there is no skill heterogeneity among workers, and firms can rely on either informal or formal labor. In this section, I first set up and solve that model, which functions as an important benchmark.

The firm takes wages as given and produces a homogenous good whose price is normalized to

[30] A similar version of this figure can be found in Gulek and Vives-i Bastida (2025).

Figure C.5: IV and SIV estimates



(a) Reduced-form

(b) 2SLS

one. The firm's objective function can be written as follows:

$$\max_{\ell_i,\ell_f} F(\ell_i,\ell_f) - \tau(\ell_i)w_i - w_f\ell_f \tag{11}$$

where $\tau(\ell_i)$ is the expected cost of hiring informal workers. In particular, I assume that $\tau(\ell_i) = \ell_i^{1+\gamma}$ with $\gamma > 0$, which satisfies the convex cost structure assumed in the literature (Ulyssea, 2018). The production function F has a CES form.

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

where θ is the Hicks-neutral productivity term, $0 < \alpha < 1$ 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 η is the share parameter of informal labor input.

Given this setup, the first-order conditions of a profit-maximizing firm are given by:

$$\begin{aligned} \theta\alpha\eta\ell_i^{\rho-1-\gamma}(\eta\ell_i^{\rho} + (1-\eta)\ell_f^{\rho})^{\frac{\alpha-\rho}{\rho}} &= w_i(1+\gamma) \\ \theta\alpha(1-\eta)\ell_f^{\rho-1}(\eta\ell_i^{\rho} + (1-\eta)\ell_f^{\rho})^{\frac{\alpha-\rho}{\rho}} &= w_f \end{aligned} \tag{12}$$

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 12.

Equilibrium

To close the model, I need to specify the labor supply. Let $L_i^{N,S}(w_i)$ and $L_f^{N,S}(w_f)$ denote the informal and formal labor supply curves of natives. As in the baseline model, labor supply curve in either sector is independent of the wage in the other sector, ruling out workers' ability to transition between informal and formal sectors.

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

$$
\begin{aligned}
L_i^S(w_i) &= L_i^D(w_i, w_f) \\
L_f^S(w_f) &= L_f^D(w_i, w_f)
\end{aligned}
\tag{13}
$$

The effects of an informal labor supply shock

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

$$
\begin{aligned}
(\rho - 1 - \gamma) log L_i &= log w_i + log(1+\gamma) - log(\alpha\eta) - \frac{\alpha-\rho}{\rho} log(\eta L_i^\rho + (1-\eta) L_f^\rho) \\
(\rho - 1) log L_f &= log w_f + log(1+\tau_w) - log(\alpha(1-\eta)) - \frac{\alpha-\rho}{\rho} log(\eta L_i^\rho + (1-\eta) L_f^\rho)
\end{aligned}
\tag{14}
$$

Fix $w_f = \overline{w_f}$ (assuming binding minimum wage), and differentiate w.r.t. w_i

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

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:

$$
\epsilon_{L_i,w_i} = -\frac{1-\rho-(\alpha-\rho)s_f}{(1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f + (1-\rho)s_i]}
\tag{16}
$$

and

$$
\epsilon_{L_f,w_i} = -\frac{(\alpha-\rho)s_i}{(1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f + (1-\rho)s_i]}
\tag{17}
$$

Equations 16 and 17 formalize two intuitive results. First, $\epsilon_{L_i,w_i} < 0$ for all potential parameter values, meaning as informal wages decrease, firms demand more informal labor. However, the effect on the formal labor demand is more nuanced. The sign of this elasticity depends solely on the sign of $\alpha - \rho$. When the labor share of production α is less than the CES parameter ρ, the elasticity of formal labor demand becomes positive, meaning formal labor demand goes down when informal wages go down.

To grasp the logic underlying this outcome, think about how the marginal productivity of a formal worker shifts upon the employment of an informal worker. 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 Q-complementarity between workers. Consequently, the firm demands more formal labor, leading to a negative elasticity of formal labor demand $\epsilon_{L_f,w_i} < 0$. However, as α decreases, hiring an additional worker incurs productivity losses for the rest of the workers due to decreasing returns. If α 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 Q-complementarity between workers. Consequently, an informal labor supply shock that reduces informal wages can incentivize firms to substitute formal workers with informal workers.[31]

It is also worth showing how the labor demand elasticity for formal workers changes with firm size. The following is the equivalent of Proposition 4 without skill heterogeneity:

Proposition 6. *The magnitude of the elasticity of formal labor demand w.r.t. informal wages is decreases in firm productivity:* $\frac{\partial|\epsilon_{L_f,w_i}|}{\partial\theta} < 0$

Proof. Notice that the productivity θ enters into equations 16 and 17 implicitly through s_i and s_f, the share of informal and formal labor in production. Therefore, I first show how productivity θ impacts these shares.

Lemma 7. *Share of formal workers increases in firm size:* $\frac{\partial s_f(\theta)}{\partial\theta} > 0$

Proof.

$$s_f = \frac{(1-\eta)l_f^{\rho}}{\eta\ell_i^{\rho} + (1-\eta)\ell_f^{\rho}}$$

$$\frac{\partial log(s_f)}{\partial log(\theta)} = \rho s_i(\epsilon_{l_f,\theta} - \epsilon_{\ell_i,\theta})$$

Taking the derivative of FOCs. w.r.t. θ, we get:

$$\epsilon_{\ell_i,\theta} = \frac{\rho-1}{\rho-1-\gamma}\epsilon_{\ell_f,\theta}$$

Since both elasticities are positive, this means that $\epsilon_{\ell_i,\theta} < \epsilon_{\ell_f,\theta}$. Therefore, we get

$$\frac{\partial log(s_f)}{\partial log(\theta)} = \rho s_i(\epsilon_{l_f,\theta} - \epsilon_{\ell_i,\theta}) > 0$$

□

Intuitively, the more productive the firm, the more it wants to produce, which it can only do by hiring more labor. Because of the increasing costs of hiring informal workers, the firm relies more on formal labor as it gets bigger: $\frac{\partial s_f(\theta)}{\partial\theta} > 0$. Hence, to see how the labor demand elasticities

[31] An alternative way to generate this qualitative prediction is presented in Delgado-Prieto (2024), 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 additional free parameters.

change when the firm gets bigger, it is sufficient to check how they change w.r.t. formal labor share s_f.

After some algebra, the following can be shown:

$$\frac{\partial \epsilon_{L_f,w_i}}{\partial s_f} = \frac{(\alpha-\rho)(1-\rho+\gamma)(1-\alpha)}{((1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f+(1-\rho)s_i])^2}$$

The denumerator is always positive, and since both α and ρ are bounded above by 1, we get:

$$Sign(\frac{\partial \epsilon_{L_f,w_i}}{\partial \theta}) = Sign(\alpha-\rho) = -Sign(\epsilon_{L_f,w_i}))$$

Naturally, it follows:

$$\frac{\partial |\epsilon_{L_f,w_i}|}{\partial \theta} < 0$$

□

D.2 Derivations of labor demand elasticities in the baseline model

Taking the first order conditions w.r.t. the three labor types, we get:

$$\begin{aligned}
[\ell_i]: &\quad \theta\eta_1 L^{\rho_1-1}(\eta_1 L^{\rho_1}+(1-\eta_1)\ell_h^{\rho_1})^{\frac{\alpha-\rho_1}{\rho_1}}\eta_2\ell_i^{\rho_2-1}L^{1-\rho_2}-(1+\gamma)\ell_i^{\gamma}w_i=0\\
[\ell_f]: &\quad \theta\eta_1 L^{\rho_1-1}(\eta_1 L^{\rho_1}+(1-\eta_1)\ell_h^{\rho_1})^{\frac{\alpha-\rho_1}{\rho_1}}(1-\eta_2)\ell_f^{\rho_2-1}L^{1-\rho_2}-w_f=0\\
[\ell_h]: &\quad \theta(1-\eta_1)\ell_h^{\rho_1-1}(\eta_1 L^{\rho_1}+(1-\eta_1)\ell_h^{\rho_1})^{\frac{\alpha-\rho_1}{\rho_1}}-w_h=0
\end{aligned} \tag{18}$$

Taking the natural logarithms of FOCS, and taking the derivative w.r.t. natural logarithm of informal wages w_i to get the following three equation system:

$$\begin{aligned}
&(\alpha-\rho)(s_L\epsilon_{L,wi}+S_h\epsilon_{H,wi})+(\rho_1-\rho_2)\epsilon_{L,wi}-(1-\rho_2+\gamma)\epsilon_{l_i,w_i}=1\\
&(\alpha-\rho)(s_L\epsilon_{L,wi}+S_h\epsilon_{H,wi})+(\rho_1-\rho_2)\epsilon_{L,wi}-(1-\rho_2)\epsilon_{l_f,w_i}=0\\
&(\alpha-\rho)(s_L\epsilon_{L,wi}+S_h\epsilon_{H,wi})+(\rho_1-1)\epsilon_{H,wi}=0
\end{aligned} \tag{19}$$

where $\epsilon_{L,wi} = s_{li,wi}\epsilon_{li,wi} + s_{lf,wi}\epsilon_{lf,wi}$. Rearranging the third line, we get our first result:

$$\epsilon_{H,wi} = \frac{(\alpha-\rho)s_L}{1-\rho_1-(\alpha-\rho_1)s_H}\epsilon_{L,wi} \tag{20}$$

This proves Proposition 3. Plugging this into first and second lines, and rewriting $\epsilon_{L,wi}$ in terms of $\epsilon_{li,wi}$ and $\epsilon_{lf,wi}$, we get:

$$\epsilon_{lf,wi} = \frac{(\alpha-\rho_2)(1-\rho_1-s_H(1-\rho_2))+(\rho_1-\rho_2)(1-\rho_2)s_H}{(1-\rho_1)(1-\rho_2-(\alpha-\rho_2)s_{lf})-(\alpha-\rho_1)(1-\rho_2)s_{li}s_H}s_{li}\epsilon_{li,wi} \tag{21}$$

Plugging equation 21 back into equation 19 we get:

$$\epsilon_{li,wi} = \frac{(1-\rho_1)(1-\rho_2-(\alpha-\rho_2)s_{lf}) - (\alpha-\rho_1)(1-\rho_2)s_{li}s_H}{(1-\rho_1)\left[(\alpha-\rho_2)((1-\rho_2)s_{li}+(1-\rho_2+\gamma)s_{lf}) - (1-\rho_2+\gamma)(1-\rho_2)\right] + s_H(1-\rho_2)s_{li}(\alpha-\rho_1)\gamma} \tag{22}$$

Given the derivations of the baseline model, I can now provide the proofs to the propositions in the main text.

First, to see that $\epsilon_{li,wi} < 0$ for all parameter values, notice that the elasticity is negative for both $s_H = 0$ and $s_H = 1$. Moreover, $Sign(\frac{\partial \epsilon_{li,wi}}{\partial S_H}) = Sign(\rho_1 - \alpha)$, hence the elasticity is always increasing or decreasing as S_H moves between 0 and 1. Therefore, it always remains negative. This proves Proposition 1.

It is important to note that equation 21 equals 17 when the labor share of high-skill labor s_H equals zero. Meaning, as the share of high-skill labor goes down, the model with skill heterogeneity naturally collapses into the model without skill heterogeneity. This is an important insight which will be beneficial in the proofs.

To prove Proposition 2, notice that the denumerator of equation 21 is positive for all values of parameters. This can be proven by plugging in $s_H = 0$ and $s_H = 1$ and seeing that denumerator is negative for both values. Given that the denumerator is linear in s_H, it is always negative. The sign of the numerator, on the hand, depends on s_H and ρ_1. When $\rho_2 > \rho_1$ and $\rho_2 > \alpha$, i.e., the elasticity of substitution between informal and formal labor is large enough, then the numerator is also negative. This can be proven via proof by contradiction.

Suppose the numerator is positive: $(\alpha - \rho_2)(1-\rho_1) - s_H(1-\rho_2)(\alpha-\rho_1) > 0$. This implies:

$$\begin{aligned}
(\alpha-\rho_2)(1\rho_1) &> s_H(1-\rho_2)(\alpha-\rho_1) \\
\Rightarrow \alpha-\rho_2 &> s_H\frac{1-\rho_2}{1-\rho_1}(\alpha-\rho_1) \quad \text{Using } \rho_2 > \alpha\text{, we get:} \\
\rho_2-\alpha &< s_H\frac{1-\rho_2}{1-\rho_1}(\rho_1-\alpha) \quad \text{Using } \rho_2 > \rho_1 \text{ and } s_H < 1\text{, we get:} \\
\rho_2 &< \rho_1 \quad \text{, a contraction}
\end{aligned}$$

This finishes the proof of Proposition 2. Proving Proposition 4 is straightforward using the following equivalence result.

Lemma 8. *When $\rho_1 = \alpha$, the labor demand elasticities of low-skill labor demand (both informal and formal) w.r.t. informal wages are numerically equivalent to their respective elasticities in the benchmark model without skill heterogeneity.*

The proof is trivial: plugging in $\rho_1 = \alpha$ into equations 22 and 21 results in equations 16 and 17, respectively. Therefore, the proof of Proposition 6 also proves Proposition 4

E Model Estimation

This section discusses the estimation of the model with firm heterogeneity. Since I do not document changes in high-skill labor demand, for estimation I focus on the model without skill heterogeneity that is introduced in Section D.1. This can be justified by assuming $\rho_1 = \alpha$ in the baseline model. In fact, the labor demand elasticities for low-skill informal and formal workers from the baseline model are numerically equivalent to the labor demand elasticities in a model without skill heterogeneity. Hence, estimating the parameters in the model without skill heterogeneity is equivalent to estimating the parameters in the model with skill heterogeneity where $\rho_1 = \alpha$ is enforced.

To analyze counterfactual policy changes, it is necessary to estimate and calibrate the four key parameters of the model: the share of labor in production α, the elasticity of substitution between the informal and formal labor $\sigma = \frac{1}{1-\rho}$, the share parameter of informal labor η, and the convex cost structure of hiring informal workers γ. The model is estimated using a minimum distance estimator. Firm heterogeneity is introduced to obtain additional moments for identification. Section E.1 sets up the full model, while Section E.2 describes the estimation method, identification, and the model's fit.

E.1 Introducing Firm heterogeneity in productivity

Building on the representative firm framework of Section D.1 I allow for firms to have different productivities denoted by $\theta \in \{\theta_1, \dots, \theta_K\}$, which enters firms' production function in a Hicks-neutral way:

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

Firm of type θ's objective function is given by:

$$\max_{\ell_i, \ell_f} F(\ell_i, \ell_f; \theta) - \ell_i^{(1+\gamma)} w_i - (1+\tau_w) w_f \ell_f$$

where τ_w is introduced to take into account that the firm pays for taxes in addition to the salaries of formal workers. The first-order conditions determine the labor demand functions of each firm of type θ:

$$\alpha \eta \ell_i^{\rho-1-\gamma} Y^{\frac{\alpha-\rho}{\alpha}} = w_i(1+\gamma)$$
$$\alpha(1-\eta)\ell_f^{\rho-1} Y^{\frac{\alpha-\rho}{\alpha}} = w_f(1+\tau_w)$$

where $Y(\theta) = \theta(\eta \ell_i^\rho + (1-\eta)\ell_f^\rho)^{\frac{\alpha}{\rho}}$ is the output produced by the firm of type θ. Solving these two equations for $L_i(\theta)$ and $L_f(\theta)$ determines the informal and formal labor demanded by firms of type θ. The total labor demand curves are given by aggregating these group-specific labor demand curves.

Given K types of firms with productivities $\theta \in \{\theta_1, \dots, \theta_K\}$, let n_j and m_j denote the ratio of informal and formal labor hired by firms of type θ_j. The aggregate informal labor demand elasticities w.r.t. informal wages are then given by weighted averages of group-specific elasticities:

$$\overline{\epsilon_{L_i,w_i}} \coloneqq \sum_{j=1}^{K} \epsilon_{L_i,w_i}(\theta_j) n_j$$
$$\overline{\epsilon_{L_f,w_i}} \coloneqq \sum_{j=1}^{K} \epsilon_{L_f,w_i}(\theta_j) m_j$$

where the group-specific labor demand elasticities are given by:

$$\epsilon_{L_i,w_i}(\theta) = -\frac{1-\rho-(\alpha-\rho)s_f(\theta)}{(1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f(\theta)+(1-\rho)s_i(\theta)]}$$
$$\epsilon_{L_f,w_i}(\theta) = -\frac{(\alpha-\rho)s_i(\theta)}{(1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f(\theta)+(1-\rho)s_i(\theta)]]}$$

where $s_i(\theta) = \frac{\eta \ell_i(\theta)^\rho}{(\eta \ell_i(\theta)^\rho + (1-\eta)\ell_f(\theta)^\rho)}$ is the share of informal labor in production for firms of type θ.

I partition the vector of parameters into two groups based on whether they are calibrated or estimated. $\alpha = 0.45$ is calibrated based on the share of labor in production in Turkey (Sevinc et al., 2021), informal wage w_i and formal wage w_f for the low-skilled are estimated using the labor force surveys, the labor tax rate is set to its statutory value $\tau_w = 0.25$. The value of τ_w corresponds to the effective tax rate for minimum wage earners. The mean formal wage for low-skill earners is inflated by 1/12 to account for the statutory severance pay rate.

E.2 Estimation Method

I take the parameters defined in the first step as given and use a Minimum Distance estimator to obtain the remaining model parameters. The model has three core parameters $\{\gamma, \eta, \rho\}$ and K productivity measures θ_K that need to be estimated. The estimator proceeds in two steps. First, it uses the model to generate the informal and formal labor demanded by each firm type. Second, it uses these inputs to compute the set of moments computed from actual data and the IV estimates. The estimate is obtained as the parameter vector that best approximates these moments.

Let $\hat{m}_N = \frac{1}{N}\sum_{i=1}^{N} m_i$ denote the vector of moments computed from data, which can include, for example, the share of informal workers hired by firms of different sizes. Let the model-generated counterpart of these moments be denoted by $m(\Phi;\Psi)$. Define $g_N(\Phi;\Psi) = \hat{m}_N - m_s(\Phi;\Psi)$; the estimator is then given by

$$\hat{\Phi} = \arg\min_{\Phi} Q(\Phi;\Psi) = \left\{ g_N(\Phi;\Psi)' W_N g_N(\Phi;\Psi) \right\} \tag{23}$$

where W_N is a positive, semi-definite weighting matrix. For simplicity, I use a diagonal matrix where each element is the inverse of the square of the empirical moment. This way, percentage deviations from the moments take equal weight.

Moments and Identification

I use nine moments from the data and my IV results to form the vector $\hat{m}_N$. HLFS asks respondents how many people work in their establishment, and group results in K categories: less than 10, between 10–24, 25–49, 50–249, and 250–499 workers. I follow this structure of the HLFS and further calculate the average number of employees in each group of firms using the census of firms in Turkey.[32] The moments I choose are (i) the size of firms in different groups (calculated using HLFS and Turkish census), (ii) the informality rate of firms in different groups (calculated using HLFS), (iii) the ratio of informal and formal labor demand elasticities (estimated in the empirical section).

This section's main goal is not to provide a rigorous proof of identification. Nonetheless, here I explain how the observed variations in data, combined with the outcomes of reduced-form analyses and the structure of the underlying model, help determine the model's parameters. In this model, the sole means by which firms can augment their output is by increasing their workforce, as labor constitutes the exclusive input in the production process. Consequently, the distinction between larger and smaller firms hinges entirely upon disparities in their productivities denoted as θ. More productive firms choose to expand their workforce. The parameter γ, which governs the marginal cost of employing informal workers, predominantly hinges on the extent to which larger firms opt for formalization at the intensive margin. For all types of firms, the share parameter η is linked to the relative productivity of formal and informal workers and, thus, is determined by the proportion of informal workers in the overall economy. The elasticity of substitution between informal and formal workers is primarily dictated by demand elasticities. For instance, the sign of the formal labor demand elasticity in isolation provides set identification for ρ as $\rho > \alpha \iff \epsilon_{L_f,w_i} > 0$. Similarly, the relative magnitudes of the elasticities of informal and formal labor demand, expressed as $\frac{\epsilon_{L_f,w_i}}{\epsilon_{L_i,w_i}} = \frac{(\alpha-\rho)s_i}{1-\rho-(\alpha-\rho)s_f}$, assist in pinpointing ρ. Holding the share of informal labor constant, this ratio exhibits a declining trend with respect to ρ.

Estimates and Model Fit

Table E.1 shows the values of all parameters. The most critical estimate is that the CES elasticity parameter ρ is 0.89, which implies an elasticity of substitution between informal and formal labor of 10. To the best of my knowledge, this is one of the first papers to estimate this elasticity. This relatively high elasticity is consistent with the Turkish context, where informal employment is often in the same sectors and even in the same firms as formal employment. It also supports the assumption of perfect substitutability between informal and formal workers in the

[32] An important detail is that I observe only formal workers in the Turkish census, whereas HLFS considers informal and formal workers combined. To account for this disparity, I first estimate the informality ratio of each group of firms using the HLFS, which I use to calculate the range of formal workers these firms should be employing on average. For example, I calculate that 58,5% of salaried workers in firms with less than 10 employees are informal, which means that these firms, on average, hire between 1–4 formal workers. I then look at the firm size distribution in the Turkish census, calculate the average formal firm size within each group, and then calculate the average total firm size by dividing by the formality rate.

Table E.1: Parameter Values

Parameter	Description	Source	Value
τ_w	Payroll tax	Statutory values	0.25
w_i	Informal wages	Calibrated	2.95
w_f	Formal wages for the low-skilled	Calibrated	4.44
α	Cobb-Douglass coefficient	Calibrated	0.54
γ	Intensive mg. cost of informal labor	Estimated	0.24
η	Informal share parameter	Estimated	0.46
ρ	CES elasticity parameter	Estimated	0.89
θ_1	Productivity of firms between 1–9 workers	Estimated	26.48
θ_2	Productivity of firms between 10–24 workers	Estimated	50.70
θ_3	Productivity of firms between 25–49 workers	Estimated	76.12
θ_4	Productivity of firms between 50–249 workers	Estimated	127.02
θ_5	Productivity of firms between 250–499 workers	Estimated	209.45
$\sigma_{i,f}$	Elasticity of substitution between informal and formal workers	Implied	9.58
ϵ_{L_i,w_i}	Average Elasticity of informal labor demand w.r.t. informal wages	Implied	-2.50
ϵ_{L_f,w_i}	Average Elasticity of formal labor demand w.r.t. informal wages	Implied	0.64
	Effect of a 1pp increase in refugee/native ratio on informal wages faced by firms	Implied	-1.32%

Note: Formal and informal hourly wage estimates are expressed as averages of log hourly earnings.

recent structural literature on the informal sector (Ulyssea, 2018, 2020).

The implied elasticity of informal and formal labor demand w.r.t informal wages are -2.50 and 0.64, respectively. The relatively large elasticity in the informal sector can be explained by the lack of institutional forces that protect workers, such as severance pay. Moreover, the model allows me to back up the decrease in informal wages faced by firms. I estimate that for every 1 pp increase in refugee/native ratio, the informal wages faced by firms decrease by 1.39%. A reduced-form test of this prediction would require observing the universe of informal wages in the economy. Unfortunately, I do not observe the wages of refugees in the HLFS, and I cannot account for the compositional change in the HLFS as it is not a panel of individuals. Instead, I use a back-of-the-envelope calculation to estimate how much the average informal wages in the economy have decreased due to the compositional effects of refugees earning less than natives. Turkish Red Crescent and WFP (2019) survey refugees in Turkey in selected regions and find that refugees earn 1058 TRY on average per month. Most of them are working informally due to the lack of work permits. Using HLFS in 2018 and restricting the data to those regions, I calculate that natives in the informal sector earn 1373 TRY on average per month. Using the 47% salaried employment rate among refugees (Turkish Red Crescent and WFP, 2019) and the 8.5% informal salaried employment rate among natives, I estimate that the average informal wage faced by firms has decrease by 1.23% just from the compositional change due to refugees. The difference between the two wage estimates may be explained by refugees' lowering wages of natives who are not displaced.

Table E.2 shows how the model performs compared to all of the targeted moments in the data. The model matches most of the moments of the data quite well. In general, there is a larger deviation between model and data in larger firms in contrast to smaller firms.

Table E.2: Model Fit

Moments	Source	Data	Model
Size of firm			
1–9 workers	HLFS and census	4.38	4.32
10–24 workers	HLFS and census	15.36	15.24
25–49 workers	HLFS and census	34.85	34.52
50–249 workers	HLFS and census	98.64	106.10
250–499 workers	HLFS and census	341.22	312.98
Share of informality			
1–9 workers	HLFS	0.59	0.58
10–24 workers	HLFS	0.29	0.28
25–49 workers	HLFS	0.16	0.16
50–249 workers	HLFS	0.071	0.079
250–499 workers	HLFS	0.043	0.038
Ratio of demand elasticities	IV estimates	-3.82	-3.89

F Alternative Identification Strategies and Contentious Findings

As described in the introduction, the literature examining the effects of Syrian refugees on Turkish labor markets has produced largely inconclusive results despite employing various identification strategies. While there exists some consensus on one finding, Del Carpio and Wagner (2015); Ceritoglu et al. (2017); Aksu et al. (2022) all document a decline in informal employment among natives following the refugee influx, other labor market outcomes remain contested. For instance, Del Carpio and Wagner (2015) report increased formal employment, but only among low-skill men, whereas Akgündüz and Torun (2020), using the same dataset, assert that high-skill employment (predominantly formal) increased. Further complicating the picture, Aksu et al. (2022) find that refugee inflows generated gender-differentiated effects: increased formal employment for men but decreased employment for women. This latter finding stands in partial contradiction to Erten and Keskin (2021), who document that women's employment rate decreases. Most recently, Cengiz and Tekgüç (2022), employing a generalized synthetic control method to address pre-trends, conclude that the refugee shock produced no employment losses among natives whatsoever. Table F.1 summarizes these studies' identification strategies, pre-trend adjustments, time periods, and key findings relevant to the present analysis.

Table F.1: Papers studying the effect of Syrian Refugees on Turkish Labor Markets

Paper	Strategy	Pre-trend adjustment	Time-period	Relevant conclusion
Del Carpio and Wagner (2015)	IV-DiD	None	2011 and 2014	Decrease in informal employment and increase in formal employment
Tumen (2016)	DiD	None	2010–2013	Decrease in informal employment and increase in formal employment
Ceritoglu et al. (2017)	DiD	None	2010–2013	Decrease in informal employment and increase in formal employment
Akgündüz et al. (2023)	IV-DiD	Aggregate region-year f.e.	2010–2015	Increase in natives' task complexity of high-skill natives
Erten and Keskin (2021)	IV-DiD	None	2006–2014	No impact among men, decrease in employment among women
Aksu et al. (2022)	IV-DiD	Aggregate region-year f.e. & linear trends using post data	2004–2015	Decrease in informal employment, increase in formal employment for men, decrease in formal employment for women
Cengiz and Tekgüç (2022)	SC	SC	2004–2015	No effect on employment of natives

This paper proposes that these divergent findings stem from two critical methodological shortcomings: (1) failure to disaggregate employment into components governed by distinct economic mechanisms, particularly salaried versus non-salaried employment, and (2) inadequate accounting for pre-trends in instrumental variable difference-in-differences (IV-DiD) designs. In what follows, I elaborate on these limitations, particularly in earlier contributions to this literature, and demonstrate how the analytical framework developed in this paper reconciles these seemingly contradictory findings.

F.1 Unaddressed Pre-trends and Endogeneity in earlier papers

The pioneering contributions to this literature (Del Carpio and Wagner, 2015; Tumen, 2016; Ceritoglu et al., 2017) suffer from a critical methodological limitation: the absence of pre-trend analysis and adjustment. This omission undermines causal inference by conflating differential economic trajectories with treatment effects. For instance, while these studies correctly identify a greater increase in low-skill native employment in southeastern Turkey during their observation periods (2010–2013 for Tumen (2016); Ceritoglu et al. (2017); 2011–2014 for Del Carpio and Wagner (2015)), they fail to recognize that this differential existed and was even more pronounced during 2004–2010, predating the refugee crisis entirely.

The methodological concerns extend beyond trajectory differences. Tumen (2016) and Ceritoglu et al. (2017) employ ordinary least squares estimation to compare regions with varying refugee presence before and after the influx. This approach fundamentally suffers from endogeneity bias, as refugees' location decisions likely correlate with unobserved regional economic conditions. While Del Carpio and Wagner (2015) advance the literature by introducing travel distance as an instrument for refugee settlement patterns, thereby addressing the endogeneity concern, their analysis still fails to account for preexisting differential trends. Therefore, both endogeneity and differences in economic trajectories result in inconsistent estimates in this earlier body of work.

F.2 Inadequate adjustment for pre-trends in the IV-DiD design

Aksu et al. (2022) is the first study to identify the pre-trends in the IV design. They employ two strategies to account for these trends: (1) controlling for linear trends in a nonsaturated IV regression and (2) controlling for aggregate region-year fixed effects. The latter strategy is later adopted by Akgündüz et al. (2023). In this subsection, I show that these strategies not only fail to mitigate bias but actually amplify it in the Turkish context.

F.2.1 The pitfalls of controlling for aggregate region-year fixed effects in Turkey

The labor force statistics in Turkey are representative at the NUTS-2 level, comprising of 26 regions. Let i denote a NUTS-2 region. The inclusion of aggregate region-year fixed effects involves defining broader regional categories $k \in K$ and incorporating interaction terms between these K regions and T time periods. To evaluate whether these additional controls effectively eliminate pre-trends, I estimate the following nonparametric event study specification:

$$y_{i,t} = \sum_{j \neq 2010} \theta_j (\text{year}_j \times Z_i) + f_k * f_t + f_i + \eta_{i,t} \tag{24}$$

where f_k is an aggregate region indicator, f_i and f_t are region and year fixed effects, respectively. Aksu et al. (2022) use two different aggregate region definitions: the 12 NUTS-1 regions defined by Turkstat and a broader 5-region categorization (termed "NUTS-0") defined by the authors.

Figure F.1 presents the results, focusing on formal salaried employment among low-skill men, a critical outcome where my findings diverge from theirs. Panel A displays estimates controlling for NUTS-0 region-year fixed effects, Panel B for NUTS-1 region-year fixed effects, and Panel C replicates my preferred specification from the main text. Crucially, region-year fixed effects not only fail to eliminate pre-trends but actually exacerbate them, as evidenced by comparing pre-2010 estimates across panels. Using the region-year fixed effects specification, a one standard deviation increase in the instrument predicts a two percentage point increase in formal salaried employment in the pre-period (2004–2010), followed by no change in the post-period (2010–2016). This pattern leads their IV-DiD design to erroneously attribute a positive effect to refugee inflows through a mechanical process: estimating null effects post-treatment, negative coefficients pre-treatment, and subtracting the latter from the former to yield positive estimates. This is unlikely to reflect a causal effect of immigration.

F.2.2 Adjusting for linear trends in a nonsaturated regression

The alternative approach employed by Aksu et al. (2022) involves controlling for linear trends within a nonsaturated regression. Specifically, after defining the inverse-distance share Z_i, they define a shift-share instrument by interacting the shares with the total refugee population in Turkey each year.

$$Z_{it} = \underbrace{H_t}_{\text{shift}} \times \underbrace{Z_i}_{\text{share}}$$

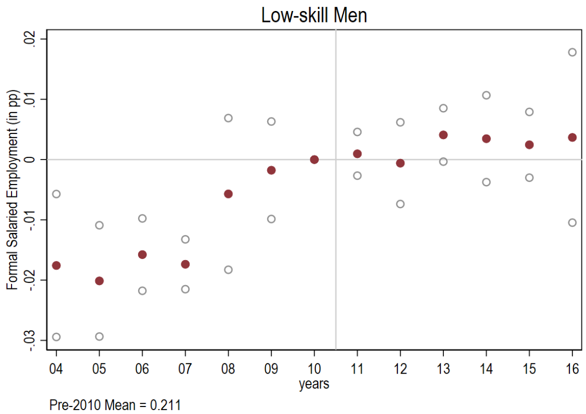
where H_t denotes the total number of refugees in year t. This shift-share instrument is then employed in the IV regression

$$\begin{aligned} y_{it} &= \beta R_{it} + f_i + f_t + f_i * t + \epsilon_{it} \\ R_{it} &= Z_{it} + g_i + g_t + g_i * t + \eta_{it} \end{aligned} \tag{25}$$

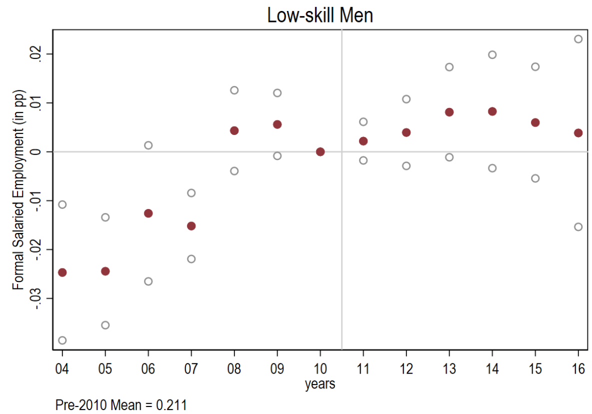
where $f_i * t$ and $g_i * t$ represent region-specific linear trends in the structural and first-stage equations.

This specification suffers from two critical flaws. First, it produces biased estimates of structural linear trends by incorporating post-treatment data, a well-documented issue in the difference-in-differences literature. As Wolfers (2006) notes: "A major difficulty in difference-in-difference analyses involves separating out preexisting trends from the dynamic effects of a policy shock. [...] This problem —that state specific trends may pick up the effects of a policy and not just pre existing

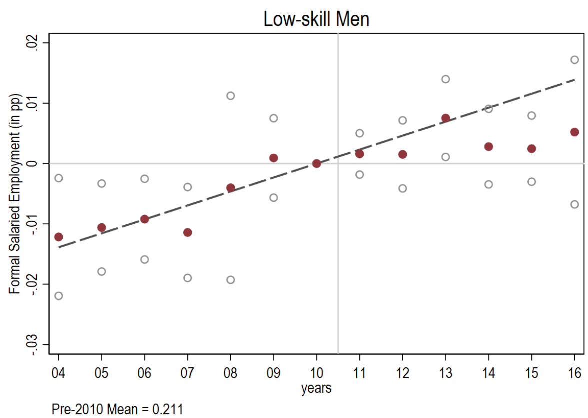
Figure F.1: Comparison of identification strategies in the literature: region*year fixed effects



(a) Nuts0-year fixed effects



(b) Nuts1-year fixed effects



(c) The preferred linear trend method

Notes: NUTS-1 categories are taken from Turkstat, NUTS-0 definitions are taken from Aksu et al. (2022). In the preferred method, the nonparametric estimates are plotted together with the linear trend that is estimated using the parametric event study design.

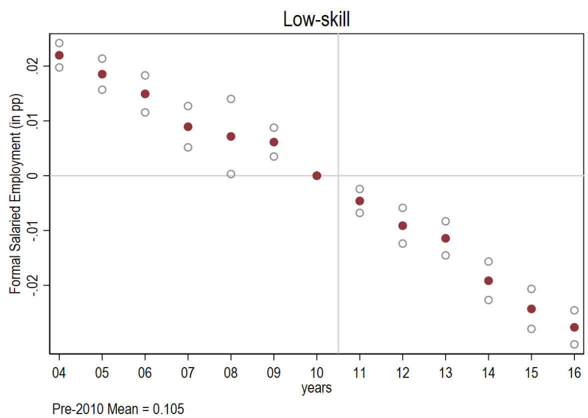
trends— is quite general." This methodological challenge motivates my approach of saturating the post-period while estimating the linear trend, which forces the linear trend to be estimated using only the pre-period variation Dobkin et al. (2018).

To provide visual evidence for this flaw, I estimate the following event study design while controlling for region-specific linear trends inside the nonsaturated regression.

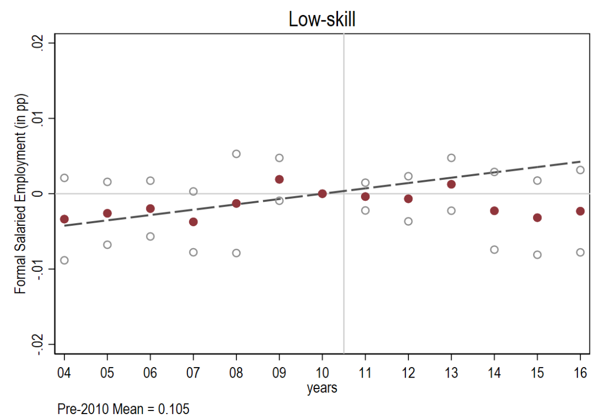
$$y_{i,t} = \sum_{j \neq 2010} \theta_j(\text{year}_j \times Z_i) + f_i * t + f_i + f_t + \eta_{i,t} \tag{26}$$

where $f_i * t$ represents the region-specific linear trends. I estimate this equation where the outcome variable is the formal employment of low-skill natives. Figure F.2 compares these estimates (Panel A) with the estimates from my preferred design (Panel B). The evidence clearly demonstrates that linear trend controls in nonsaturated regressions amplify rather than reduce bias.

Figure F.2: Comparison of identification strategies in the literature: linear trend



(a) Nonsaturated linear trends



(b) The preferred design

Notes: In Panel A, a different linear trend is estimated for each NUTS-2 region. In the preferred method, the nonparametric estimates are plotted together with the linear trend that is estimated using the parametric event study design.

The second flaw emerges in the first-stage estimation. Since treatment intensity is zero before 2011 and monotonically increases thereafter, the linear trend in the first stage necessarily adopts a positive slope, creating a *pseudo treatment*; a change in predicted treatment intensity before actual treatment begins. Figure F.3 contrasts nonparametric first-stage estimates from the nonsaturated model (Panel A) with my preferred specification (Panel B), revealing that the former erroneously generates pre-treatment first-stage effects while the latter does not.

In conclusion, while Akgündüz et al. (2023) and Aksu et al. (2022) acknowledge pre-trends in the IV design and attempt adjustments through linear trend controls and aggregate region-year fixed effects, neither approach adequately addresses these trends. In fact, both methods systematically amplify bias from differential trajectories, undermining causal inference in this context.

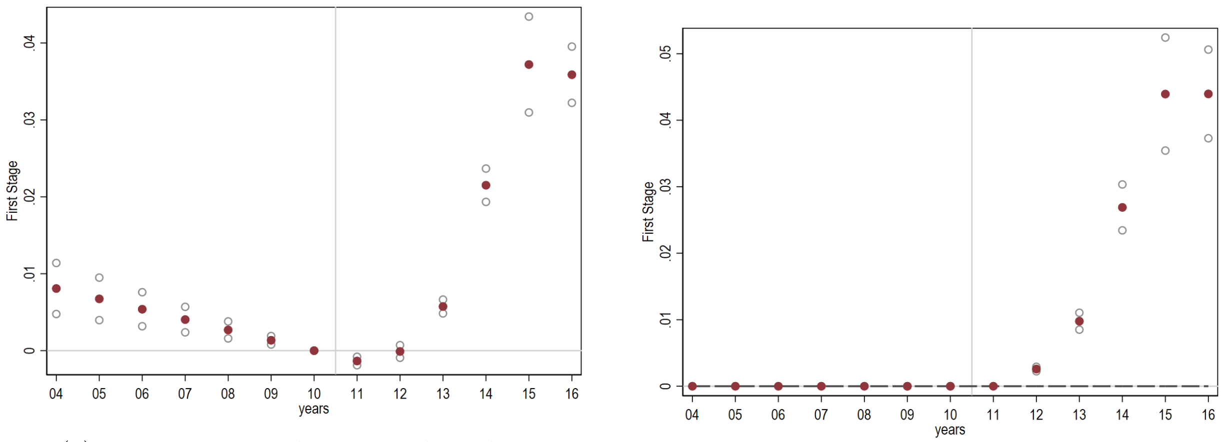
F.3 Different economics behind salaried and non-salaried work hides the dis-employment effects

There are two papers in the literature that provide alternative identification strategies. Erten and Keskin (2021) control for region-time varying controls, and Cengiz and Tekgüç (2022) apply a generalized synthetic control estimator. Here, I briefly discuss their methodologies.

F.4 Using "bad controls" to address pre-trends

Erten and Keskin (2021) conclude that Syrian refugee inflows reduce employment among native women but not men, a finding they leverage as a first-stage in examining the second-stage effects on

Figure F.3: Comparison of identification strategies in the literature: first stage estimates



(a) First stage with linear trends within nonsaturated regression

(b) First stage in the preferred design

Notes: In Panel A, a different linear trend is estimated for each NUTS-2 region. In the preferred method, the nonparametric estimates are plotted together with the linear trend that is estimated using the parametric event study design.

intimate partner violence. Their core argument posits that since the refugee shock selectively impacted women's labor market opportunities while sparing men's, it constitutes a quasi-experimental setting for investigating how women's economic prospects influence intimate partner violence.

Their empirical strategy employs an individual-level specification:

$$Y_{ipt} = \beta R_{pt} + \gamma X_{ipt} + \sigma Z_{pt} + \delta_p + \delta_t + \epsilon_{ipt}$$

where Y_{ipt} is the outcome for individual i in province p in year t, X_{ipt} includes individual-level controls including education, age and mother tongue; Z_{pt} comprises time-varying province controls that include the volume of trade of each province with Syria and the volume of trade at the beginning of the period interacted with time dummies (both in logs). They instrument regional treatment R_{pt} with the standard distance-based shift-share instrument similar to the one used in the present manuscript.

This specification suffers from three critical methodological flaws. First, controlling for trade volumes constitutes a "bad control" (Angrist and Kugler, 2003) if refugee inflows influence trade patterns. Second, standard models of trade and labor markets suggest that total trade exposure, rather than bilateral exchange with a relatively minor trading partner like Syria, should drive labor market effects. Third, incorporating 26 region-specific log-trade volumes interacted with time indicators introduces an opaque identification strategy where the precise variation driving treatment effect estimates becomes ambiguous.

Beyond these econometric concerns, the fundamental divergence between their findings and mine stems from an economic misspecification. As Figure 8 clearly shows, both low-skill men and women experience comparable losses in salaried employment after refugee inflows. However, men respond to these losses by transitioning into non-salaried positions, which leaves their aggregate employment rate unchanged. This reallocation pattern, visible only by disaggregating employment by job type, reconciles the apparent contradictions in the literature and underscores the importance of distinguishing between employment categories governed by different economic mechanisms.

F.5 Addressing pre-trends without resolving endogeneity

Cengiz and Tekgüç (2022) employ the generalized synthetic control (GSC) method developed by Xu (2017), an extension of the synthetic control approach (Abadie et al., 2010), which constructs synthetic control regions that are "similar" to the host regions receiving immigrants. Their analysis finds no significant adverse employment effects for native Turkish workers, which they attribute to immigrants' positive impact on local demand.

This methodological approach requires careful examination. The immigration literature has traditionally relied on instrumental variables to address the fundamental endogeneity concern that migrants select destinations based on local labor market conditions. This selection may reflect both observable economic trajectories and unobservable contemporaneous shocks. While IV strategies effectively mitigate selection bias, in the present setting, they fail to resolve differential trends issues. Conversely, synthetic control methods successfully address differential trends but leave the endogeneity problem unresolved. Even when host and non-host regions exhibit parallel pre-trends, the synthetic control estimator would produce results comparable to OLS, leaving labor economists justifiably concerned about migrants' location decisions and their potential correlation with unobserved local economic shocks.

A methodological advance that integrates these approaches, the Synthetic IV (SIV) developed by Gulek and Vives-i Bastida (2025) after this paper's initial draft, offers a promising solution. Section C.3 details this methodology, which simultaneously leverages IV to address endogeneity and synthetic controls to account for unobserved confounders. Importantly, Figure C.5 confirms that my central findings remain robust when implementing SIV.

Notwithstanding potential endogeneity concerns, Cengiz and Tekgüç (2022)'s primary conclusion, the absence of statistically significant disemployment effects for Turkish natives, stems largely from overlooking the transition of low-skill men to non-salaried positions after losing their salaried jobs. As Figure 8 illustrates, both men and women experience salaried job losses, but men transition to non-salaried positions at a near one-to-one rate. This pattern produces three simultaneous effects: men's overall employment rate remains stable, women's employment rate declines, and because men's employment share exceeds women's, the aggregate employment rate appears unchanged. Combined with the positive bias inherent in their non-IV approach, Cengiz and Tekgüç (2022)'s null result and their explanation invoking demand-side offsets are unsurprising.

Had Cengiz and Tekgüç (2022) conducted disaggregated analyses by employment type or gender, even their GSC strategy would likely have detected native displacement effects. However, the previously undocumented phenomenon of men's transition to non-salaried work after losing salaried positions, a novel contribution of my analysis, has prevented previous studies from isolating the relevant economic mechanisms operating in the Turkish context.

F.6 Resolving contradictions in the literature

The effects of Syrian refugees on Turkish labor markets have been extensively studied, generating a body of literature with seemingly irreconcilable findings. Although previous studies have made valuable contributions to the literature, little effort has been directed toward explaining the underlying reasons for these contradictory results. This detailed methodological examination demonstrates how my findings not only differ from existing work but also help to reconcile these apparent inconsistencies.

Two fundamental challenges have complicated the analysis of Syrian immigration's labor market impacts in Turkey. First, the southeastern regions most exposed to refugee inflows were following distinct economic trajectories compared to the rest of Turkey, causing early studies to misattribute differential trends to causal effects (Del Carpio and Wagner, 2015; Tumen, 2016; Ceritoglu et al., 2017). Although Aksu et al. (2022) and Akgündüz et al. (2023) identified these pre-trends, their adjustment strategies inadvertently magnified rather than mitigated the resulting bias. Second, displaced low-skill men systematically transitioned into non-salaried employment, rendering aggregate "employment rates" an inadequate indicator of true labor market outcomes. This labor market reallocation explains why Cengiz and Tekgüç (2022) found null effects in the aggregate while Erten and Keskin (2021) observed gender-differentiated impacts: both studies did not recognize this crucial substitution mechanism driving their results.

A key contribution of this paper lies in isolating and accounting for these two confounding forces, thereby providing a coherent framework that explains the literature's conflicting findings. After implementing appropriate econometric adjustments and separating employment into salaried and non-salaried types, the evidence clearly demonstrates that low-skill Syrian immigrants who could only work in the informal sector displaced low-skill natives, both men and women, from salaried positions across both formal and informal sectors. This cross-sector displacement indicates high substitutability between formal and informal labor in production. The gender-differentiated response to this displacement is striking: men predominantly shifted into non-salaried work while women disproportionately exited the labor force or remained unemployed. Furthermore, the evidence suggests an extensive margin effect in which marginally productive firms strategically remained unregistered to access informal labor more easily.

G Additional Empirical Results for the Online Appendix

G.1 Tables showing hetetogeneity analyses across sex, formality, industry and employment type

Table G.1: Refugees' effect on the employment rate of natives

Formality	Sex	All	Agriculture	Manufacturing: Textiles	Manufacturing: Other	Construction	Services: Market	Services: Non-market
		(1)	(2)	(3)	(4)	(5)	(7)	(8)
	Pooled	-0.277	-0.218	-0.0349**	0.00670	-0.0581***	-0.0185	0.0453
		(0.250)	(0.189)	(0.0162)	(0.0103)	(0.0204)	(0.0504)	(0.0281)
Informal	Men	-0.298	-0.173	-0.0334***	0.0144	-0.125***	-0.0206	0.0398*
		(0.201)	(0.131)	(0.0106)	(0.0176)	(0.0418)	(0.0856)	(0.0208)
	Women	-0.257	-0.252	-0.0356	-0.000296	0.00219	-0.0210	0.0502
		(0.315)	(0.255)	(0.0267)	(0.00691)	(0.00145)	(0.0199)	(0.0455)
	Pooled	-0.00660	0.0531	-0.0556**	-0.000736	-0.0294**	0.0363	-0.0103
		(0.0644)	(0.0447)	(0.0236)	(0.0251)	(0.0116)	(0.0299)	(0.0136)
Formal Low-skill	Men	0.0528	0.0993	-0.104**	0.0243	-0.0661**	0.114*	-0.0150
		(0.117)	(0.0863)	(0.0410)	(0.0492)	(0.0270)	(0.0610)	(0.0289)
	Women	-0.0914**	0.00890	-0.0224	-0.0279*	-0.00212	-0.0389***	-0.00898
		(0.0432)	(0.0187)	(0.0168)	(0.0150)	(0.00204)	(0.0148)	(0.00760)
	Pooled	0.000265	0.0117	-0.135***	0.0582	-0.0461	0.0441	0.0675
		(0.164)	(0.0245)	(0.0357)	(0.0526)	(0.0316)	(0.0473)	(0.138)
Formal High-skill	Men	0.176	0.0335	-0.183***	0.135*	-0.0679	0.156**	0.102
		(0.200)	(0.0393)	(0.0397)	(0.0749)	(0.0539)	(0.0643)	(0.148)
	Women	-0.00729	-0.00711	-0.0202	0.00226	0.00512	-0.0241	0.0368
		(0.171)	(0.00623)	(0.0281)	(0.0252)	(0.0120)	(0.0451)	(0.180)

The 2SLS estimates come from estimating equation 5 using natives' informal, low-skill formal, and high-skill formal salaried employment rates. Standard errors are clustered at the region level. Industry codes are determined according to ISIC. Details can be found following this link: https://ilostat.ilo.org/resources/concepts-and-definitions/classification-economic-activities/ Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table G.2: Refugees' effect on the salaried employment rate of natives

Formality	Sex	All	Agriculture	Manufacturing Textiles	Manufacturing Other	Construction	Services Market	Services Non-market
		(1)	(2)	(3)	(4)	(5)	(7)	(8)
Informal	Pooled	-0.169***	-0.0611**	-0.0307**	-0.00358	-0.0641***	-0.0307*	0.0207
		(0.0538)	(0.0273)	(0.0150)	(0.00979)	(0.0211)	(0.0162)	(0.0190)
	Men	-0.298***	-0.0694**	-0.0417***	-0.00351	-0.137***	-0.0388	-0.00792
		(0.0826)	(0.0291)	(0.0156)	(0.0197)	(0.0433)	(0.0243)	(0.0223)
	Women	-0.0536	-0.0540	-0.0202	-0.00347	0.00240**	-0.0250*	0.0467
		(0.0829)	(0.0352)	(0.0174)	(0.00497)	(0.00106)	(0.0132)	(0.0412)
Formal Low-skill	Pooled	-0.128***	0.00281	-0.0668***	-0.0185	-0.0261**	-0.0121	-0.00700
		(0.0381)	(0.00674)	(0.0237)	(0.0260)	(0.0127)	(0.0184)	(0.0126)
	Men	-0.190**	0.0116	-0.130***	-0.0139	-0.0597**	0.00926	-0.00723
		(0.0792)	(0.0121)	(0.0420)	(0.0500)	(0.0294)	(0.0360)	(0.0251)
	Women	-0.104**	-0.00514	-0.0214	-0.0285*	-0.00109	-0.0390**	-0.00888
		(0.0416)	(0.00404)	(0.0159)	(0.0148)	(0.00166)	(0.0157)	(0.00797)
Formal High-skill	Pooled	-0.125	-0.0131	-0.139***	0.0191	-0.0391	-0.00835	0.0551
		(0.130)	(0.00972)	(0.0399)	(0.0487)	(0.0337)	(0.0354)	(0.130)
	Men	-0.0700	-0.0168	-0.188***	0.0614	-0.0600	0.0589	0.0747
		(0.134)	(0.0147)	(0.0454)	(0.0658)	(0.0565)	(0.0408)	(0.142)
	Women	-0.0131	-0.00319	-0.0220	0.00940	0.00666	-0.0490	0.0450
		(0.177)	(0.00293)	(0.0265)	(0.0272)	(0.0122)	(0.0382)	(0.179)

The 2SLS estimates come from estimating equation 5 using natives' informal, low-skill formal, and high-skill formal salaried employment rates. Standard errors are clustered at the region level. Industry codes are determined according to ISIC. Details can be found following this link: https://ilostat.ilo.org/resources/concepts-and-definitions/classification-economic-activities/ Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table G.3: Refugees' effect on the non-salaried employment rate of natives

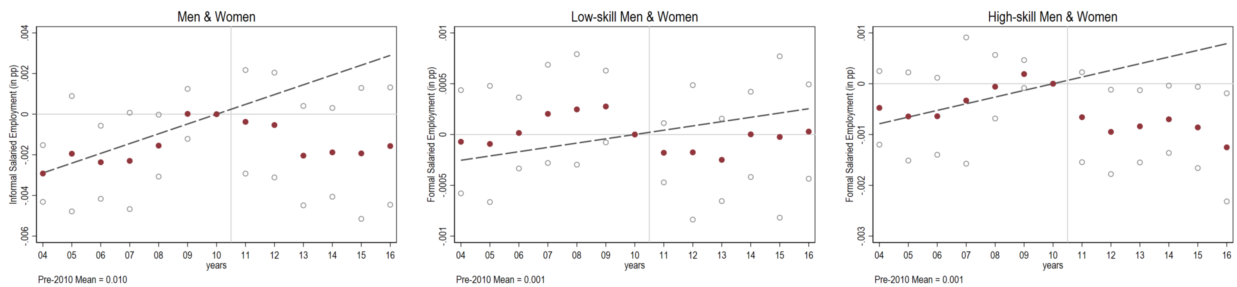
Formality	Sex	All	Agriculture	Manufacturing: Textiles	Manufacturing: Other	Construction	Services: Market	Services: Non-market
		(1)	(2)	(3)	(4)	(5)	(7)	(8)
Informal	Pooled	-0.108	-0.156	-0.00423	0.0103	0.00597	0.0122	0.0246**
		(0.220)	(0.167)	(0.0112)	(0.00822)	(0.00611)	(0.0382)	(0.0109)
	Men	0.000352	-0.104	0.00824	0.0179	0.0119	0.0183	0.0477**
		(0.192)	(0.111)	(0.00684)	(0.0161)	(0.0120)	(0.0717)	(0.0198)
	Women	-0.203	-0.198	-0.0155	0.00318	-0.000210	0.00397	0.00355
		(0.258)	(0.240)	(0.0189)	(0.00323)	(0.000546)	(0.0114)	(0.00820)
FormalLow-skill	Pooled	0.121**	0.0503	0.0113**	0.0177**	-0.00334	0.0484***	-0.00327
		(0.0493)	(0.0439)	(0.00440)	(0.00696)	(0.00378)	(0.0187)	(0.00776)
	Men	0.242**	0.0876	0.0260***	0.0382**	-0.00637	0.105***	-0.00775
		(0.0968)	(0.0828)	(0.00963)	(0.0150)	(0.00813)	(0.0401)	(0.0137)
	Women	0.0125	0.0140	-0.000991	0.000528	-0.00102	0.0000680	-0.0000968
		(0.0230)	(0.0208)	(0.00186)	(0.00162)	(0.000623)	(0.00454)	(0.00361)
FormalHigh-skill	Pooled	0.126**	0.0247	0.00383	0.0391***	-0.00698	0.0525**	0.0123
		(0.0566)	(0.0177)	(0.00732)	(0.0106)	(0.00802)	(0.0257)	(0.0108)
	Men	0.246**	0.0503*	0.00566	0.0735***	-0.00797	0.0971**	0.0270
		(0.102)	(0.0292)	(0.0154)	(0.0206)	(0.0121)	(0.0444)	(0.0168)
	Women	0.00579	-0.00393	0.00177	-0.00714	-0.00155	0.0249	-0.00826
		(0.0292)	(0.00549)	(0.00452)	(0.00572)	(0.00179)	(0.0182)	(0.00838)

The 2SLS estimates come from estimating equation 5 using natives' informal, low-skill formal, and high-skill formal salaried employment rates. Standard errors are clustered at the region level. Industry codes are determined according to ISIC. Details can be found following this link: https://ilostat.ilo.org/resources/concepts-and-definitions/classification-economic-activities/ Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

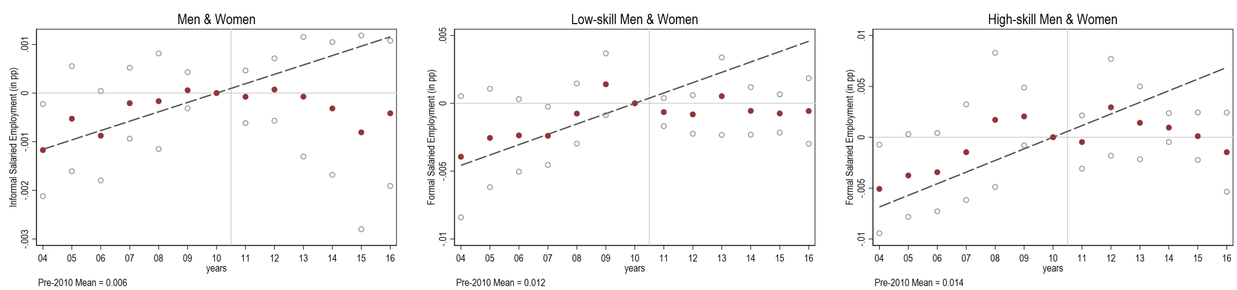
G.2 Event study figures of the 2SLS estimates

Figure G.1: Event study figures of industry specific estimates in Figure B.1

(a) Agriculture



(b) Textile



(c) Other Manufacturing

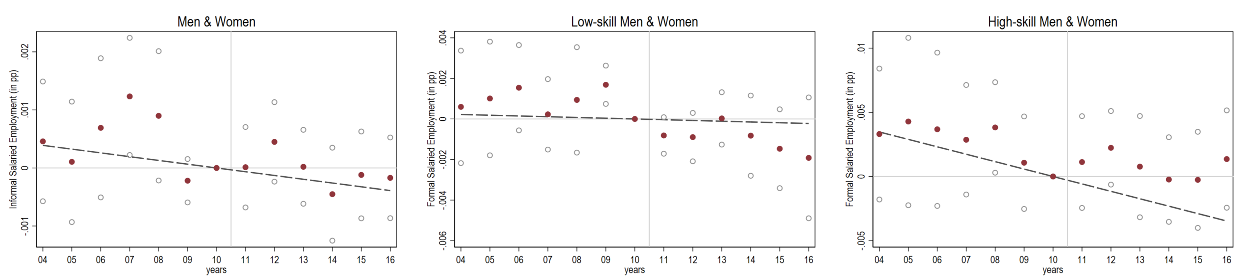
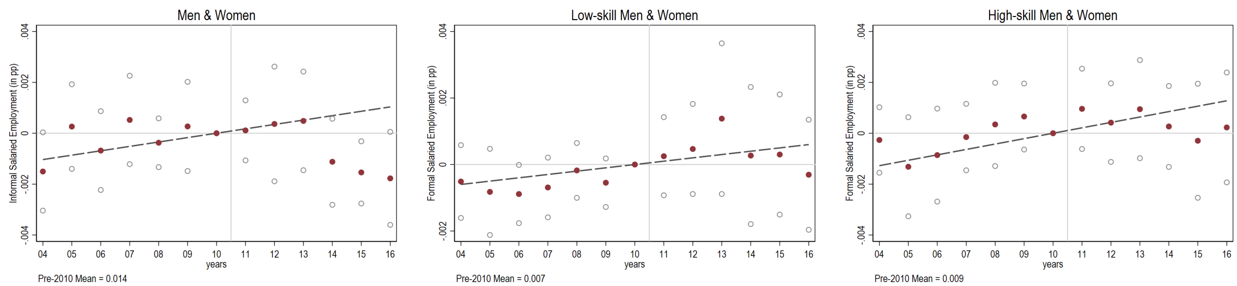
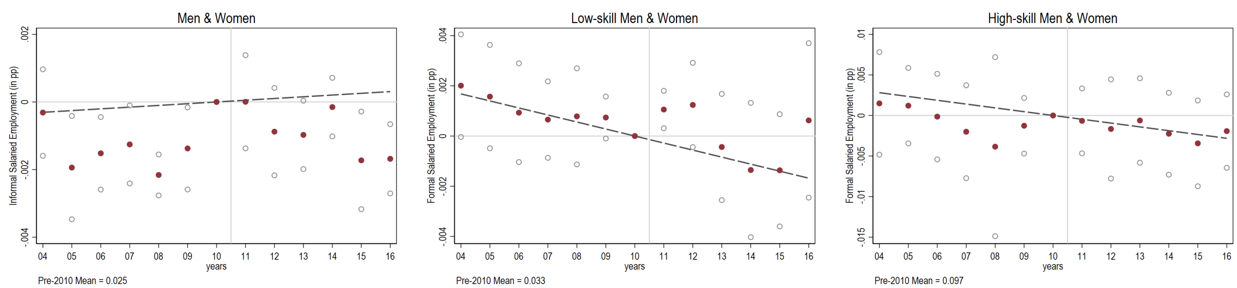


Figure G.1: Event study figures of industry specific estimates in Figure B.1 (cont.)

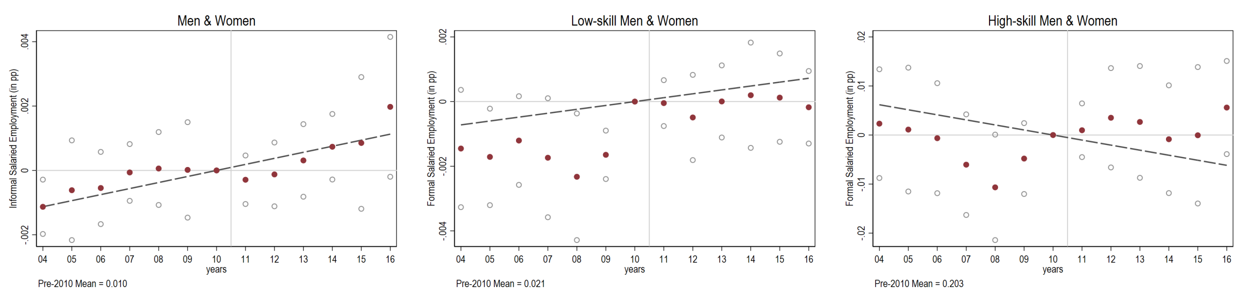
(d) Construction



(e) Market Services

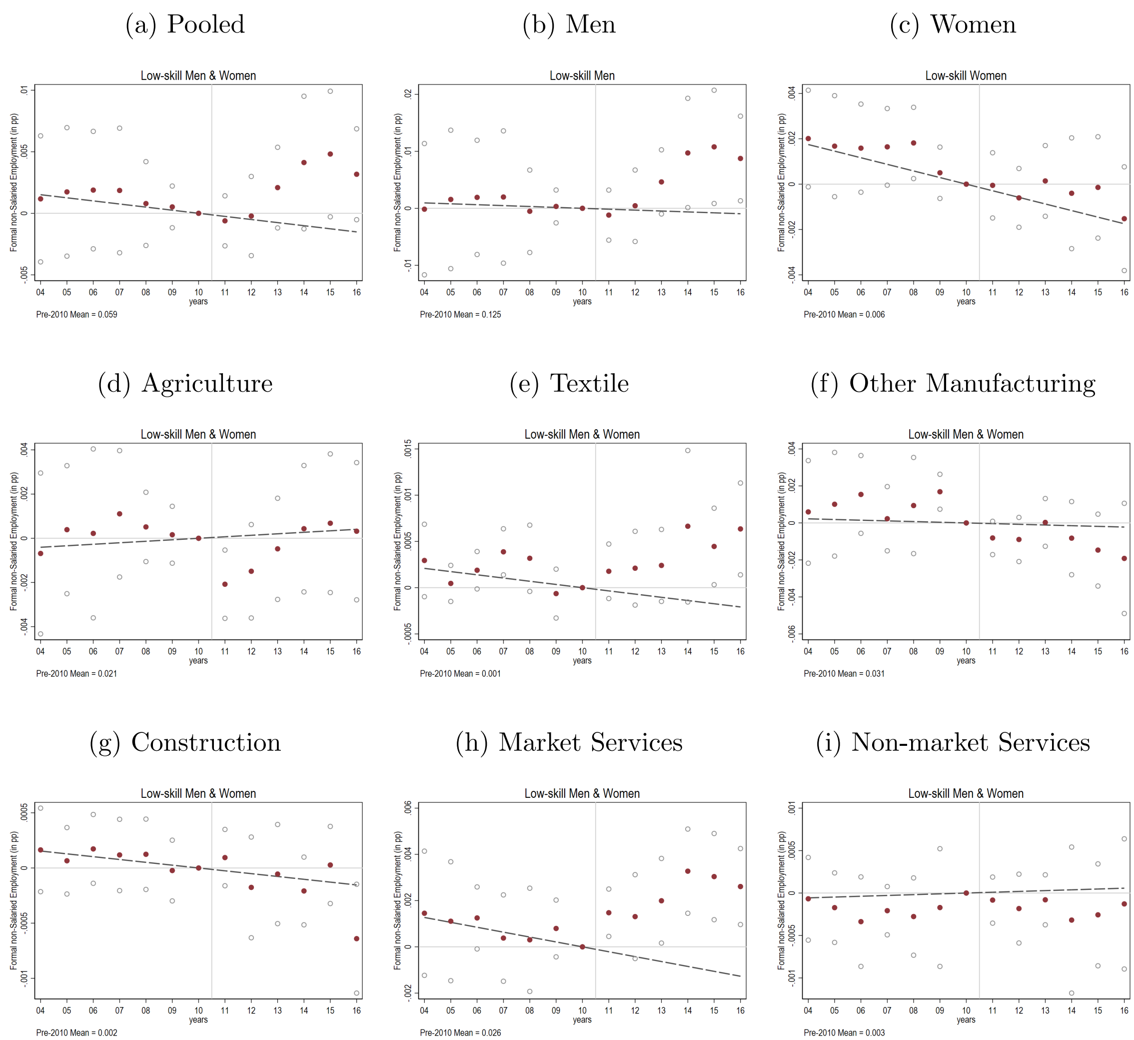


(f) Non-market Services



Notes: The points in each figure represent the estimated effects of event time shown in equation 3. The hollow circles present the 95 percent confidence intervals. The dashed line represents the estimated pre-2010 linear relationship between outcome and instrument * event time from the parametric event study in equation 4 with the level normalized to match the nonparametric estimates.

Figure G.2: Event study figures of the estimates in Figure 8
Impact on the formal non-salaried employment of low-skill natives



Notes: The points in each figure represent the estimated effects of event time shown in equation 3. The hollow circles present the 95 percent confidence intervals. The dashed line represents the estimated pre-2010 linear relationship between outcome and instrument * event time from the parametric event study in equation 4 with the level normalized to match the nonparametric estimates.