Effects of Immigrants on Non-host Regions Evidence from the Syrian Refugees in Turkey

Ahmet Gulek and Tishara Garg*

November 30, 2024

Most recent draft here.

Abstract

This paper investigates how immigration-induced wage shocks can propagate beyond the regions receiving immigrants through the production network. Using the Syrian refugee crisis in Turkey as a quasi-experiment and the near universe of domestic firm-to-firm transaction data from VAT records, we show that the immigration shock propagates both forward and backward along the supply chain. Firms in non-host regions who directly or indirectly buy from host regions demand more labor. Firms who sell to host regions weakly increase their sales. Estimates imply an elasticity of substitution between labor and intermediate goods of 0.76 and an elasticity of substitution of nearly 1 between intermediates. Counterfactual analyses show that the spillover effects on non-host regions are economically meaningful when the host regions are central nodes of the domestic trade network. For example, a 1% increase in labor supply in Istanbul decreases real wages in Istanbul by 0.56% and increases real wages in the average non-host city by 0.38%.

Keywords: Immigration, production network, trade spillovers

*Job market paper for Ahmet Gulek: PhD student in Economics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA (e-mail: agulek@mit.edu). Ahmet Gulek is specially indebted to Daron Acemoglu, Josh Angrist, and Amy Finkelstein. We are also thankful to David Atkin, David Autor, Arnaud Costinot, Dave Donaldson, Adam Solomon, Edward Wiles, Emma Wiles, participants in the MIT Labor/PF Seminar, Labor Lunch and Trade Lunch. I acknowledge support from the George and Obie Shultz Fund at MIT.

1 Introduction

The last decade has seen a quadrupling of refugees globally, from 11 million in 2012 to 46 million today.[1] During this period, Turkey has received 3.6 million Syrian refugees, which has increased the labor supply of several Turkish provinces by up to 82%. Such a large increase in labor supply in host regions is likely to change prices of goods, which can induce general equilibrium effects throughout the economy. Therefore, the labor market consequences of this massive labor supply shock for the Turkish economy depend on the magnitude of these general equilibrium effects.

There are three key economic mechanisms by which an immigration shock propagates through the supply network to impact labor demand. First, immigrants reduce the wages and therefore the prices charged by firms in the host region. This reduction in prices propagates forward along the supply chain to firms who directly or indirectly buy from the host region. Whether these "upstream exposed" firms increase or decrease their labor demand is governed by the substitutability between labor and intermediate goods. The second and third forces capture the demand spillovers from backwards propagation, which we label "downstream exposure" effects. If intermediate goods are gross substitutes, then firms whose production costs fall more sharply (i.e., those who are upstream exposed) gain market share. Consequently, they demand more from their suppliers, who observe an increase in sales. Furthermore, when intermediate goods are more substitutable with other intermediates than with labor, immigrant-intensive firms increase their demand for intermediate goods, which creates a positive demand spillover for their suppliers. Together, these three economic forces shape the labor market effects of immigrants across the economy.

In this paper, we present theoretical analysis formalizing these three forces, empirical evidence testing for their existence, and counterfactual exercises that quantitatively examine the impact of immigration on real wages and welfare across regions.

Our model has two main ingredients. First, firms use a CES aggregate of local labor and an intermediate good in production, which itself is a CES aggregate of goods from other firms in all regions. Second, firms charge prices based on an exogenously given markup. Therefore, they pass along reductions in production costs, via change in prices of labor and goods, to their prices. In this framework, we characterize the impact of immigration on real wages in host and non-host regions as a function of the shape of the production network and two elasticities of substitution: one between labor and intermediate goods and another between different intermediate goods. Therefore, estimating these two elasticities, together with knowledge of the baseline input-output network, is sufficient to calculate the general

[1] Author's calculations using data from UNHCR. Appendix Figure B.2 provides more details.

equilibrium effects of immigration on labor demand in all regions.

We estimate these two elasticities by studying the effects of Syrian immigrants on firms in non-host regions in Turkey. Using the near universe of firm-to-firm transactions from VAT records, we calculate the model-defined upstream and downstream exposures to immigration for all formal firms in Turkey. We instrument for the regional immigration treatment with a shift-share IV, where the shift is the aggregate number of Syrian refugees in Turkey in a given year, and the share is the relative travel distance from the Syrian border. Specifically, the instrument captures quasi-random variation in immigration intensity across regions and years. This regional immigration shock then turns into firm-level upstream and downstream exposure treatments through firms' input-output matrices at baseline. We further relax the share-exogeneity assumption embedded in our design (Goldsmith-Pinkham et al., 2020) by applying Synthetic IV (Gulek and Vives-i Bastida, 2024).[2]

By comparing firms in the same region-industry cells who are differentially exposed to immigrants through their baseline trading partners, we obtain three findings. First, firms who directly or indirectly buy from host regions increase their labor demand: they hire more workers, they increase their payroll and their spending on labor as a share in their production costs. This implies that labor and intermediate goods are gross complements, with an estimated elasticity of substitution of 0.76. Second, firms who sell to upstream exposed firms do not see a change in their sales. This shows that firms do not substitute between different suppliers as a share of their spending, which implies an elasticity of substitution between intermediate goods of 1. Third, large firms who sell to the host regions observe a positive, yet noisy increase in their sales. This implies that intermediate goods are more substitutable between each other than with labor, which is consistent with our first two findings. These results remain similar in a series of robustness checks of the identification strategy.

Our empirical results demonstrate the existence of trade spillovers of immigration. To quantify the magnitude of these spillovers, we calibrate the model using our VAT data and the estimates of structural elasticity parameters. We simulate the actual treatment and calculate the real wage changes throughout the economy. Surprisingly, we find that the immigrant-to-native ratio in any region predicts the changes in real wages almost perfectly. This shows that the general equilibrium effects are mostly negligible in the aggregate. Whereas some trade occurs between the host south-east regions and the rest of Turkey, which enables us to estimate the structural elasticity parameters, the amount of interregional trade is too low to have meaningful effects.

[2]Note that the actual shares are functions of both firm-level input-output matrices and regional shares of the shift-share IV.

To investigate whether this result is generalizable across regions and skill-intensity of immigrants, we conduct counterfactual analyses. In our main counterfactual exercise, we treat each of the 81 provinces of Turkey separately with a 1% increase in labor supply and calculate the real wage change in host and the average non-host region. In 76 out of 81 potential host regions, because of the absence of spillovers, the real wages decrease by around 1% in the host region and increase by less than 0.02% in non-host regions. In contrast, a 1% increase in labor supply in Istanbul decreases real wages in Istanbul by 0.56% and spills over to increase the real wage in the average non-host region by 0.38%. Further analyses show that, while the population and economic development of host cities are correlated with spillovers, centrality within the production network is a better predictor. For example, Kocaeli and Antalya have similar populations and GDP, yet immigration to Kocaeli creates eight times larger spillovers than immigration to Antalya because the former is a more central node in the domestic trade network.

Other counterfactual exercises, where we fix the *number* of immigrants and let the immigrant-to-native ratio vary across regions, provide a similar result: the same number of immigrants creates larger spillovers in central nodes of the trade network. In terms of welfare, immigration to central nodes creates an order of magnitude larger welfare gains. The intuition is that immigration reduces production costs and prices, which helps consumers. When they settle in central nodes, more regions in the economy benefit from these cost reductions. Lastly, comparing the welfare gains between low-skill and high-skill immigration shows that high-skill immigrants create larger spillovers because high-skill-intensive industries trade more across regions than low-skill-intensive industries.

Our paper contributes to the extensive empirical literature studying the economic effects of immigration (seminal papers include Card (1990, 2001); Borjas (2003); Ottaviano and Peri (2012)).[3] Despite 30 years of work, whether immigrants lower natives' wages is still debated (Borjas, 2017; Peri and Yasenov, 2019). Our main contribution to this literature is showing, both theoretically and empirically, that the effects of immigrants spread through the supply chain via general equilibrium effects. These spillover effects on the non-host regions are economically significant when the host regions are central nodes in the domestic trade network. In such settings, comparing host to non-host regions, as is often done in the immigration literature, does not estimate the effect of immigration. In the Turkish setting, it would have overestimated the decline in real wages if immigrants had settled at central nodes. Our model shows that such a design could also underestimate immigrants' effect on wages in other settings depending on the technology parameters of the economy.

We also contribute to a branch of the immigration literature that focuses on refugee

[3]See Hanson (2009); Lewis and Peri (2015); Dustmann et al. (2016) for reviews of the literature.

crises (Hunt, 1992; Friedberg, 2001; Borjas and Monras, 2017). Several papers investigated the impact of the refugee crises of the last decade on host countries' labor markets and found larger displacement effects on natives than what is typically found in the literature on immigration.[4] One differentiating factor between refugee crises and economic migration that can explain this discrepancy based on our findings is that refugees settle in regions closer to the border they arrive from, which are typically not the most economically developed regions, whereas most voluntary immigration in the world occurs toward larger cities. Our results show that interregional trade can flatten the labor demand curve and compress the decrease in real wages in the host region, which explains the discrepancy between the labor market effects of refugee crises and voluntary immigration episodes.

A related literature investigates how the effects of immigration interact with output tradability (Dustmann and Glitz, 2015) and international trade (Caliendo et al., 2021; Brinatti, 2024). The paper that is closest to ours, Burstein et al. (2020) formalize how industry tradability shapes local labor market adjustments to immigration. We build on their findings by showing how the production network shapes local labor market adjustments to immigration. In addition to industry tradability, who industries, buy from and sell to matter first-order in our analysis.

We also contribute to a growing literature studying the propagation of technology and factor shocks through the production network. Acemoglu et al. (2012, 2016b, 2017), Baqaee and Farhi (2019) explore the conditions that enable microeconomic shocks to spread through input-output networks, leading to aggregate fluctuations.[5] Empirically, previous work has shown how trade shocks (Acemoglu et al., 2016a) and natural disasters (Barrot and Sauvagnat, 2016; Boehm et al., 2019; Carvalho et al., 2021) propagate along the supply chain. The closest empirical work to ours is Akgündüz et al. (2024), who examine the propagation of the Syrian immigration shock through the supply chain in Turkey. They show positive spillovers on firms' sales and employment through first-degree trade linkages in host regions. We extend their analysis by formalizing the mechanisms through which immigration impacts firms in non-host regions, testing these empirically, quantifying general equilibrium effects, and demonstrating when such spillovers are significant at the aggregate level.

The paper is organized as follows. Section 2 introduces the data and institutional background. Section 3 develops the model and isolates the economic forces by which an immigration induced wage shock to a region can spread through the production network to other regions. Section 4 presents the empirical results. Section 5 concludes.

[4] See Gulek (2024) for the Syrian refugee crisis in Turkey and Bahar et al. (2024) for the Venezuelan refugee crisis in Colombia.

[5] See Carvalho (2014); Carvalho and Tahbaz-Salehi (2019) for a review of the literature on production networks.

2 Background and Data

2.1 Syrian Refugee Crisis in Turkey

The Syrian Civil War started in March 2011. By 2017, 6 million Syrians had sought shelter outside of Syria, primarily in the neighboring countries Turkey, Lebanon, Jordan, and Iraq. With 3.6 million registered Syrian refugees, Turkey hosts the highest number of refugees in the world. Figure 1a shows how the number of Syrian refugees in Turkey has evolved over time. Their numbers remained small until the end of 2012 but increased substantially after. Turkey hosted around 170 thousand refugees by 2012, 500 thousand by 2013, 1.6 million by 2014, 2.5 million by 2015, and around 3.6 million by 2019.[6]

The Turkish government initially tried to host the Syrians in refugee camps in the south-eastern part of the country across the Turkish-Syrian border. However, these camps quickly exceeded capacity as the number of arriving refugees increased. The refugees thus dispersed across Turkey in heterogeneous quantities.[7] Figure 1c shows the distribution of the number of Syrian refugees per 100 natives in Turkey at the province level.[8] 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 the identification strategy.

Syrian refugees are significantly less educated than Turkish natives.[9] Figure 1b compares the education levels of Syrian refugees in Turkey with those of Turkish natives. For example, 21% of Syrian refugees did not complete primary school, compared to 12% of Turkish natives. Additionally, 83% of Syrian refugees do not have a high school diploma, in contrast to 61% of Turkish natives. Given the potential for educational downgrading (Dustmann et al., 2013) and the fact that most Syrian refugees have only basic proficiency in Turkish (Crescent and Programme, 2019), the influx of Syrian refugees can be interpreted as a low-skill labor supply shock to the Turkish labor market.

Most Syrians in Turkey do not have formal labor market access, which further limits the types of firms and industries they can work at. As of March 2019, only 31,000 Syrian refugees (1.5% of the working-age Syrians) had work permits. This feature of the immigration shock does not limit the generalizability of the present paper's findings. Gulek (2024) shows that

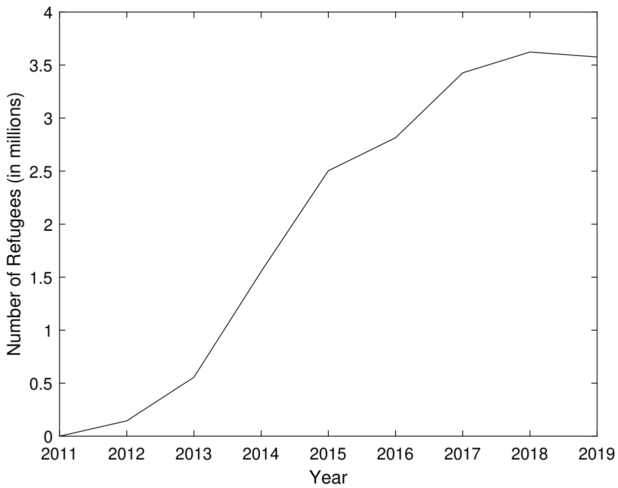
[6]The number of refugees in Turkey across years and provinces are acquired from the Directorate General of Migration Management of Turkey.

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

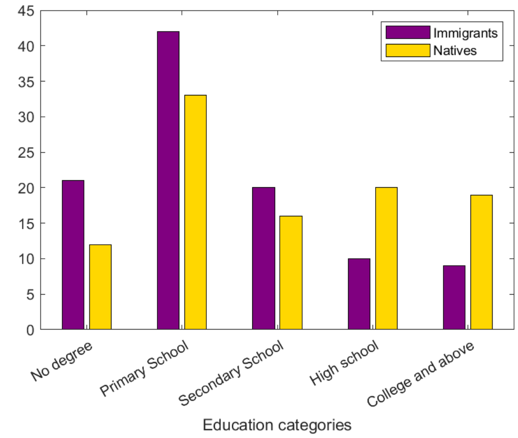
[8]Turkey does not share the education and age break-down of refugees at the province level, which prevents the empirical investigation from exploiting that variation.

[9]This is due to two main factors. Firstly, Syria was less developed than Turkey, resulting in a less educated workforce. Secondly, highly educated Syrians were more likely to seek refuge in Europe.

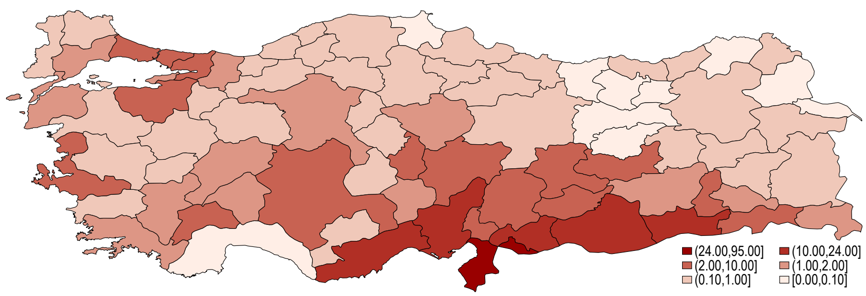
Figure 1: Statistics on the Syrian Refugees in Turkey



(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 2019

Source: Data on the number of Syrian refugees in a given year and province comoes from Directorate Generale of Migration Management of Turkey. Data on the educational attainment of refugees come from surveys on ESSN recipients. Data on natives' educational attainments come from the household labor force surveys conducted by Turkstat.

informal and formal labor in Turkey are highly substitutable in production. This implies that the informal immigration shock lowers wages in both the informal and formal sectors.

The Turkish government initially tried to host the Syrians in refugee camps in the south-eastern part of the country across the Turkish-Syrian border. However, these camps quickly exceeded capacity as the number of arriving refugees increased. The refugees thus dispersed across Turkey in heterogeneous quantities.[10] Figure 1c shows the distribution of the number of Syrian refugees per 100 natives in Turkey at the province level.[11] Refugees are more densely located in regions closer to the border, and with historically more Arabic speakers.

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

[11] Turkey does not share the education and age break-down of refugees at the province level, which prevents the empirical investigation from exploiting that variation.

These two facts constitute the one of the building blocks of our identification strategy.

2.2 Data

We integrate five datasets covering all formal firms in Turkey between 2006–2019. The Ministry of Industry and Technology maintains all the datasets and uses the same firm identifier, which enables us to merge them. Our analysis focuses on the manufacturing sector unless otherwise specified.

These datasets are as follows. First, the value-added tax (VAT) data report the value of all domestic firm-to-firm trade that exceeds 5,000 Turkish liras (about \$3,333 in 2010) in a given month. Second, from the income statements, we use the yearly gross sales of each firm. Third, from the firm registry, we extract each firm's province and two-digit industry code according to the Nomenclature Statistique des Activités Économiques dans la Communauté Européenne (NACE), the standard industry classification in the European Union. Fourth, from the customs data, we use information on total exports by firm. Fifth, from the employer-employee data, we collect the average number of workers, total labor costs and average wages per worker per each year.

We also rely on labor force surveys conducted by the Turkish statistical institute the calculate natives' employment rates and the skill intensity of 2-digit industries. These surveys are representative at NUTS-2 level.

Appendix Section A provides the details and the summary statistics about the data.

3 Theory

In this section we formalize how a decrease in wages due to immigration in one region can spillover to other regions through the production network, and develop structural equations that directly map to our reduce-form results.

3.1 Setup

The economy consists of N firms indexed by i, R regions indexed by r, where each region is endowed with L_r labor.[12] Each firm operates in one region: r_i denotes the region of firm i. Firms use intermediate goods and local labor in production, and sell their output as both an intermediate good to other producers in all regions and as a final good to local consumers.

[12] Labor is assumed to be homogeneous in the baseline model, which we later relax to become a CES aggregate of labor with different skill levels.

Producers

Firm i chooses labor L_i and intermediate goods $\{x_{i,j}\}_{j=1}^n$ to minimize costs subject to a constant returns nested-CES technology

$$\min_{\{x_{ij}\}_{j=1}^n, L_i} \sum_{j=1}^{n} p_j x_{ij} + w_{r_i} L_i \quad \text{subject to}$$
$$A_i(\eta_i m_i^{\frac{\sigma_u - 1}{\sigma_i}} + (1-\eta_i) L_i^{\frac{\sigma_u - 1}{\sigma_u}})^{\frac{\sigma_u}{\sigma_u - 1}} \geq y_i$$
$$m_i = \left(\sum_{j=1}^{n} \alpha_{ij} x_{ij}^{\frac{\sigma_l - 1}{\sigma_l}} \right)^{\frac{\sigma_l}{\sigma_l - 1}}$$

where A_i is a Hicks-neutral productivity shifter, y_i is total output, p_j is the price of good j, L_i is labor used by firm i, w_r is the wage in region r, m_i is the intermediate good used by the firm, which itself is a CES bundle of goods from different firms. x_{ij} denotes how much firm i uses firm j's goods in production, where firm j can be in any region. We assume common elasticities of substitution in both the upper and lower nests: σ_u denotes the elasticity of substitution between labor and intermediate goods, and σ_l is the elasticity of substitution between different intermediate goods.[13] Constant returns to technology requires $\sum_j \alpha_{i,j} = 1$. Let C_i denote the unit cost of firm i. We assume that firms have constant and exogenous markup μ_i, and therefore set price $p_i = \mu_i C_i$.

Final Demand

All final goods consumption as well as the ownership of firms is local. We assume a representative consumer in each region r, who optimizes her Cobb-Douglas utility subject to budget constraint that equates her spending on final goods with her labor income plus (regional) firm profits.

$$\max_{\{c_{r,i}\}} \Pi_{i \in r} c_{r,i}^{\beta_i} \quad s.t. \quad \sum_{i \in r} p_i x_{0,i} = w_r L_r + \sum_{i \in r} \pi_i$$

where $c_{r,i}$ is how much the representative agent r consumes firm i's goods, and $\sum_{i \in r} \beta_i = 1$.

Labor Supply

Labor is inelasticly supplied in each region, is immobile across regions and perfectly mobile across firms in a region. This simplifying assumption shuts down spillovers across regions in

[13]The common elasticity of substitution assumption across firms can easily be relaxed. However, each different parameter comes with additional data requirements for estimation, and therefore we maintain a common elasticity assumption for empirical reasons.

labor supply.[14]

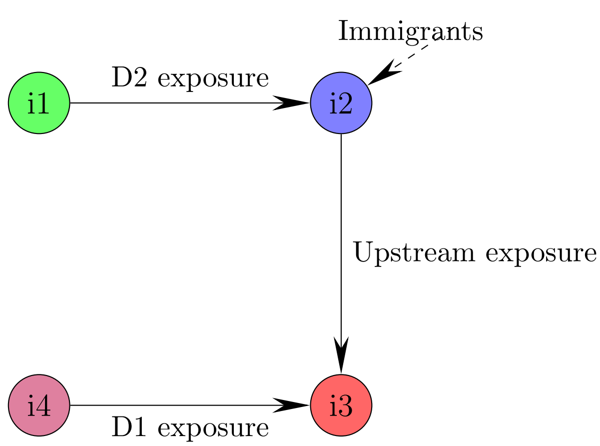
General Equilibrium

Given exogenous productivities A_i and markups μ_i, equilibrium is a set of prices p_i, wages w_r, intermediate good choices $x_{i,j}$, labor input choices l_i, outputs y_i, and final demands $c_{r,i}$, such that each producer minimizes its costs subject to technology constraints and charges the relevant markup on its marginal cost; consumers maximize their utility subject to their budget constraint, and the markets for all goods and labor clear.

3.2 Three General Equilibrium Forces

The solution to this model is notation heavy and therefore hard to follow. To facilitate exposition, we describe the three relevant economic forces here. Figure 2 depicts a simple production network with four firms in four different regions. Firm i_1 sells to i_2, and both i_2 and i_4 sell to i_3. Suppose i_2's region receives immigrants. This increase in labor supply lowers the wages and therefore the production costs of firm i_2. As firms have constant markups, lower production costs mean lower prices. This creates a chain reaction along the supply chain that goes both forward and backward.

Figure 2: Spillover Effects of Immigration Along the Input-Output Network



Notes: This figure depicts a simple input-output network where firm i_1 sells to i_2, and both i_2 and i_4 sell to i_3. Immigrant arrival to firm i_2 creates a chain reaction that impacts all other firms in this network.

First, firm i_3 benefits from immigration as the price of input from firm i_2 decreases. As i_3 faces lower input prices, it can increase or decrease its local labor demand depending on

[14] Gulek (2024) shows that changes in in- and out-migration in response to Syrian immigration has been minimal in Turkey

the substitutability between intermediate goods and labor. If labor and intermediates are gross complements, than the reduction in input prices would cause firm i_3 to increase its labor demand. We name this as the effect of "upstream exposure" of immigration, upstream because from i_3's perspective, the shock comes from upstream.

Second, the demand for i_4's goods may increase or decrease depending on the substitutability between different intermediate goods. Notice that i_2 and i_4 both supply to i_3. If intermediate goods are largely substitutable, then as i_2's prices go down compared to i_4, i_3 would demand less from i_4. As the product demand for i_4 shrinks, it reduces its labor demand. In contrast, if intermediate goods are gross complements, the opposite would take place: i_3 would increase its demand of i_4's goods, which would increase i_4's demand for local labor.

Notice that the effects on both i_3 and i_4 are part of the forward propagation channel of the immigration shock. The difference is that, while i_3 is impacted through its suppliers and therefore is upstream-exposed, i_4 is impacted through its customers and hence is downstream-exposed.

Third, the demand for i_1's goods also changes. Notice that i_2 incurs two effects: the prices of labor goes down compared to its input from i_1, and it incurs a demand shock based on i_3's choice among goods from i_2 and i_4. The former is governed by the elasticity of substitution between labor and intermediates, and the latter is governed by the elasticity of substitution between intermediates. As we shall see after having solved the model, the effect on $D1$'s sales depends on the relative magnitudes of these two elasticities. If intermediate goods are more substitutable among each other than with labor, then i_2 demands more from i_1, which increases i_1's labor demand. We call this the second downstream exposure effect, which we denote shortly as $D2$ for the rest of the paper. This captures the backward propagation of the immigration shock.

Figure 2 only depicts the first-degree trade exposures: that is, firms being impacted from their immediate customers and suppliers. However, these forces expand beyond the first-degree linkages. Firms that indirectly buy from immigrant-intensive firms (e.g., i_2) are also upstream exposed. Firms that indirectly sell to any upstream exposed firm are also $D1$ exposed. Firms who indirectly sell to immigrant-intensive firms are also $D2$ exposed. Moreover, in more complicated input-output networks, firms can have U, D1, and D2 exposures simultaneously. To understand exactly how much each firm is upstream and downstream exposed to immigrants, we need to solve the model. In return, this will help guide the empirical analysis in the next section.

3.3 Input-Output definitions

To formally understand how a labor supply shock in one region impacts the labor demand in all regions, we introduce some input-output notation and definitions.[15] Our results are comparative statics describing how the labor payments in any host and non-host region change when a host region receives immigrants. We now define accounting objects such as input-output matrices, Leontief inverse matrices, and Domar weights. These quantities have a revenue-based version and a cost-based version, and we present both. All these objects are defined at the initial equilibrium. Without loss of generality, we normalize the nominal GDP to 1. Finally, in our analytical results and counterfactuals, we assume constant markups and technology.[16]

3.3.1 Final Expenditure Shares

Let b denote the $R \times N$ matrix whose (ri)th element is equal to the share of good i in the budget of the final consumer in region r

$$b_{ri} = \frac{p_i c_i}{\sum_{j \in r} p_j c_j}$$

Let χ denote the $R \times 1$ vector of regional income shares

$$\chi_r = \frac{\sum_{j \in r} p_j c_j}{\sum_{r'=1}^{R} \sum_{j \in r'} p_j c_j}$$

where the sum of final expenditures $\sum_{r'=1}^{R} \sum_{j \in r'} p_j c_j$ is nominal GDP

3.3.2 Input-Output Matrices

To streamline the exposition, we treat labor as special endowment producer that does not use any input to produce. We form an $(N + R) \times 1$ vector of producers, where the first N elements correspond to the producers and the last R elements to the labor in each region. For labor, we interchangeably use the notation w_r or p_{N+r} to denote its wage and the notation L_{ir} or $x_{i(N+r)}$ to denote its use by firm i. The revenue-based input-output matrix Ω is the $(N + R) \times (N + R)$ matrix whose (ij)th element is equal to firm i's expenditure on inputs

[15] In particular, we follow Baqaee and Farhi (2019)'s notation closely. We deviate from their notation only when our models' regional labor markets, which is not present in their model, requires us to do so.

[16] This decision is driven primarily by the lack of data on prices. Otherwise, the model easily incorporates changes in technology and markups. For more details, see Baqaee and Farhi (2019).

from firm j as a share of its total revenues

$$\Omega_{ij} = \frac{p_j x_{ij}}{p_i y_i}$$

The first N rows and columns of Ω correspond to goods, and the last R rows and columns correspond to labor. Since labor requires no inputs, the last R rows of Ω are zeros.

The cost-based input-output matrix $\tilde{\Omega}$ is the $(N+R) \times (N+R)$ matrix whose (ij)th element is equal to i's expenditure on inputs from j as a share of its total costs

$$\tilde{\Omega}_{ij} = \frac{p_j x_{ij}}{\sum_{k=1}^{N+R} p_k x_{ik}}$$

The revenue-based and cost-based input-output matrices are related by

$$\tilde{\Omega} = diag(\mu)\Omega$$

where μ is the vector of markups, and $diag(\mu)$ is the diagonal matrix with ith diagonal element equal to μ_i.

As labor and intermediate goods appear as the sole two inputs in the upper nest of the CES production function, defining the labor share and intermediate goods share of costs is useful for exposition. We define the share of labor and intermediate good expenditures of firm i as:

$$\tilde{\Omega}_{i,L} = \frac{w_r L_i}{\sum_{k=1}^{N} p_k x_{ik} + w_r L_i} \quad ; \quad \tilde{\Omega}_{i,M} = 1 - \tilde{\Omega}_{i,L}$$

3.3.3 Leontief Inverse Matrices

We define the revenue-based and cost-based Leontief inverse matrices as

$$\Psi = (I - \Omega)^{-1} = I + \Omega + \Omega^2 + \dots, \quad \text{and} \quad \tilde{\Psi} = (I - \tilde{\Omega}) = I + \tilde{\Omega} + \tilde{\Omega}^2 + \dots$$

While the input-output matrices Ω and $\tilde{\Omega}$ capture the direct exposures of one firm to another, the Leontief inverse matrices Ψ and $\tilde{\Psi}$ capture the total exposures, direct and indirect, through the production network.

Note that the revenue-based Leontief inverse matrix Ψ encodes the backward propagation of demand, whereas the cost-based Leontief inverse matrix $\tilde{\Psi}$ encodes the forward propagation of costs.

3.3.4 Domar Weights

The revenue-based Domar weight λ_i of producer i is its sales as a fraction of nominal GDP:

$$\lambda_i \equiv \frac{p_i y_i}{nGDP} = p_i y_i$$

Similarly, the revenue-based Domar weight λ_r for labor in region r is its total labor payments $w_r L_r$.

Before stating our results, we introduce the following input-output covariance operator:

$$Cov_{\tilde{\Omega}^{(j)}}(d\ln p, \Psi_{(k)}) = \sum_i \tilde{\Omega}_{ji} d\ln p_{(i)} \Psi_{ik} - \left(\sum_i \tilde{\Omega}_{ji} d\ln p_i\right)\left(\sum_i \tilde{\Omega}_{ji} \Psi_{ik}\right)$$

where $\tilde{\Omega}^{(j)}$ corresponds to the jth row of $\tilde{\Omega}$, $d\ln p$ is the vector of price changes of all inputs, and $\Psi_{(k)}$ is the kth column of Ψ. Because the rows of $\tilde{\Omega}$ always sum up to 1 for each firm j, we can formally think of this as a covariance. It answers the question: "Among the suppliers of firm j, are the ones who decrease their prices more rely on firm i more or less for intermediate goods?" If the answer is more, the covariance term is negative.

3.4 Effects of a Labor Supply Shock on labor income

Before characterizing the full set of equilibrium changes in prices and quantities, we first build intuition as to how an immigration shock in a host region can impact the labor payments in any region. To achieve this, we take the change in prices $d\ln p$ and $d\ln w$ as given, and describe how the demand for labor and for goods change in response to these changes in prices.

Note that the labor income in region r is the sum of labor payments by all firms in that region.

$$\lambda_r = w_r L_r = \sum_{i \in r} \lambda_i \Omega_{i,L}$$

Hence, the change in labor payments is determined by the change in sales and the change in labor share of sales

$$d\ln \lambda_r = \sum_{i \in r} \frac{\lambda_i \Omega_{iL}}{\lambda_r}(d\ln \lambda_i + d\ln \Omega_{iL})$$

Therefore, to understand the impact of immigration on labor payments in all regions, we need to determine the impact on firms' sales share in GDP and labor share in sales. Propositions 1 and 2 characterize these effects.

Proposition 1. *In response to an immigration-induced wage shock, the following equation describes the change in the labor share of production costs*

$$d\ln\tilde{\Omega}_{i,L} = (1-\sigma_u)(d\ln w_{r_i} - \sum_{j=1}^{n}\frac{\tilde{\Omega}_{ij}}{\tilde{\Omega}_{iM}}d\ln p_j) \tag{1}$$

All proofs are in the Appendix.

Equation 1 captures the forward propagation of cost shocks, which is the upstream exposure effect we introduced in Figure 2. Firms' labor share is determined by the trade-off firms face between hiring labor and using intermediate goods in production. Suppose the local wages go down less than the prices of the suppliers of firm i. If labor and intermediate goods are gross complements, $\sigma_u < 1$, then the firm would increase its labor share in production.

Note that, absent changes in markups, $d\ln\tilde{\Omega}_{i,L} = d\ln\Omega_{i,L}$. Therefore, equation 1 also describes the change in the labor share of sales.

Proposition 2. *In response to an immigration-induced wage shock, the following equation describes the change in the Domar weights / sales share of firms*

$$\begin{aligned} d\ln\lambda_i = & \sum_{j=1}^{n}(1-\sigma_l)\frac{\lambda_j}{\lambda_i\mu_j}Cov_{\tilde{\Omega}^{(j)}}\left(d\ln p, \Psi_{(i)}\right) \\ & + (\sigma_u - \sigma_l)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\tilde{\Omega}_{j,l}\left(d\ln w_{r_j} - \sum_{k=1}^{n}\frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}}d\ln p_k\right)(\Psi_{ji} - I_{ji}) \\ & +\eta_i \end{aligned} \tag{2}$$

where I is the identity matrix, and $\eta_i = \frac{1}{\lambda_i}\sum_j\sum_r b_{rj}\Psi_{ji}\chi_r\left(\left(\sum_{i\in r}\frac{\pi_i}{\chi_r}d\ln\lambda_i\right) + \frac{\lambda_r}{\chi_r}d\ln\lambda_r\right)$ captures the demand spillovers of immigrants' demanding locally produced goods.

The first term captures the first downstream exposure effect: demand spillovers from firms substituting across intermediates. The immigration shock propagates forward and lowers costs throughout the supply chain. When different intermediate goods are largely substitutable, $\sigma_l > 1$, those who observe larger decreases in costs gain market share and demand more goods from their suppliers. This is captured by the covariance term, which is negative when those that observe larger decreases in costs among the suppliers of firm j are also more dependent on firm i for production. Summing across all firms in the economy and their suppliers determines the total demand spillover from substitution among intermediates.

The second term captures the second downstream exposure effect: the demand spillovers from firms substituting between intermediate goods and labor. Assume $\sigma_l > \sigma_u$, that is, the different intermediate goods are more substitutable than intermediate goods and labor. In this case, if firm j observes larger decreases in local wages than the prices of its intermediate goods, $\left(d\ln w_{r_j} - \sum_{k=1}^{n} \frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}} d\ln p_k\right) < 0$, then it will spend a larger share of its production costs on intermediate goods. This, in turn, increases the demand for firm i to the extend that firm j relies on firm i's goods, which is captured by Ψ_{ji}. Summing over all such firms determines the total demand spillover from substitution between intermediates and labor.

The third term captures the demand spillovers from changing income shares of the regions due to immigration. Immigrants increase the consumer base in the host regions. Firms that sell goods to these host regions directly or indirectly also observe an increase in their demand.[17]

Given the intuition we developed in Propositions 1 and 2, we now move on to fully characterizing the change in equilibrium prices and quantities with respect to an immigration shock $d\ln L$. Proposition 3 characterizes the change in prices of firm i as a function of changes in wages.

Proposition 3. *In response to an immigration-induced wage shock, the following equation describes the change in prices charged by firms*

$$d\ln p_i = \sum_{j=1}^{n} \tilde{\Psi}_{ij}\tilde{\Omega}_{jL} d\ln w_{r_j} \tag{3}$$

Proposition 3 shows an intuitive result. As firms have constant markups, any change in their production costs are fully represented in their prices. $\tilde{\Psi}_{ij}$ captures how much firm i depends on goods of firm j for production. $\tilde{\Omega}_{jL} d\ln w_{r_j}$ captures the change in production costs of firm j from the change in local wages. Multiplying the two terms and summing across all firm j's give us how much the production cost, and hence the price, of firm i changes in response to changes in wages.

Lastly, note that the share of labor in GDP is simply the wage times the quantity of labor in that region: $\lambda_r = L_r w_r$. Combining this with Propositions 1, 2, 3, we can fully characterize the impact of immigration on this economy

Theorem 1. *The following linear system fully describes the change in equilibrium prices*

[17] In practice, immigrants and natives can demand different type of goods. Unfortunately, the lack of data on the consumption basket of Syrian immigrants in Turkey prevents us from exploring this dimension in detail without strong assumptions. Hence, in the empirical section we assume that this force enters the error term and is not correlated with our instrument.

and quantities in response to an immigration shock $d\ln L$.

$$\begin{aligned}
d\ln\lambda_r =& \sum_{i\in r}\frac{\lambda_i\Omega_{iL}}{\lambda_r}(d\ln\lambda_i + d\ln\Omega_{iL}) \\
d\ln\Omega_{i,L} =& (1-\sigma_u)(d\ln w_{r_i} - \sum_{j=1}^{n}\frac{\tilde{\Omega}_{ij}}{\tilde{\Omega}_{iM}}d\ln p_j) \\
d\ln\lambda_i =& (1-\sigma_l)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i\mu_j}Cov_{\tilde{\Omega}^{(j)}}\left(d\ln p, \Psi_{(i)}\right) \\
&+ (\sigma_u-\sigma_l)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\tilde{\Omega}_{j,L}\left(d\ln w_{r_j} - \sum_{k=1}^{n}\frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}}d\ln p_k\right)(\Psi_{ji} - I_{ji}) \\
&+ \frac{1}{\lambda_i}\sum_j\sum_r b_{rj}\Psi_{ji}\chi_r d\ln\chi_r \\
d\ln\chi_r =& \left(\sum_{i\in r}\frac{\pi_i}{\chi_r}d\ln\lambda_i\right) + \frac{\lambda_r}{\chi_r}d\ln\lambda_r \\
d\ln p_i =& \sum_{j=1}^{n}\tilde{\Psi}_{ij}\tilde{\Omega}_{jL}d\ln w_{r_j} \\
d\ln w_r =& d\ln\lambda_r - d\ln L_r
\end{aligned} \tag{4}$$

Equation 4 presents the economic forces we have described in one system of linear equations. An important observation for the rest of the paper is that, except for the elasticity parameters σ_u and σ_l, we observe all the parameters in this equation in our pre-shock data. Therefore, estimating these two elasticities using the immigration shock is sufficient to quantify the total impact of immigration on all host and non-host regions.

Overall, the model characterizes exactly how an immigration shock propagates forward and backward and impact firms' labor demand and sales across the economy. However, it comes with certain limitations that we discuss here.

First, we assume constant structural elasticities across firms, region, and industries. In practice, this need not hold. The model can be expanded to incorporate heterogeneity in these elasticities in a straightforward way. However, it is not relevant for our setting as, we show in the next section, elasticities seem to be similar across the 2 digit industry groups within manufacturing firms. Therefore, assuming a constant elasticity of substitution is an innocuous simplifying assumption for our setting.

Second, we assume that labor does not move across regions. This simplifies the model and allows us to focus on trade spillovers. In practice, native migration responses to immigration can be an important channel by which regional labor markets can equilibrate over time

(Monras, 2020). However, in the Turkish setting, native migration responses to immigration have been limited (Gulek, 2024). Appendix Figure C.10 shows that Syrian refugees have caused only economically and statistically insignificant changes in in-migration and out-migration rates in the host regions. Therefore, this assumption is consistent with the Turkish data.

Third, one caveat of Theorem 1 is that it does not provide a simple sufficient statistic to understand when the spillover effects are small or large. This makes it hard to provide an economic intuition as to when general equilibrium effects of immigration would be meaningfully different from partial equilibrium effects. Therefore, we answer that question through counterfactuals in Section 4.7.

4 Empirical Analysis

This section presents the trade spillover effects of immigration on firms in non-host regions. We first use Propositions 1 and 2 to define the three treatments from trade exposure. The causal effects of these three treatments on firms' labor demand and sales help identify the structural elasticity parameters: the elasticity of substitution between labor and intermediates and the elasticity of substitution between different intermediate goods. We then use these elasticity parameters together with our VAT data to quantify the total effects of immigration on host and non-host regions. This enables us to quantify when the spillover effects of immigration are economically meaningful and what these spillovers imply about our understanding of the effects of immigration on labor markets.

4.1 Treatment Definitions

The model highlights three economic forces that help determine the equilibrium effects of immigration on host and non-host regions: a forward propagation of costs and demand spillovers from substitution effects. We refer to these three forces as upstream exposure effect U, the first and second downstream exposure effects $D1$ and $D2$.

Specifically, The upstream exposure of firm i at time t

$$U_{it} = \sum_{r=1}^{R} \tilde{\Psi}_{i,r} \delta_{rt} \tag{5}$$

summarizes how the regional arrival of Syrian immigrants δ_{rt} are transmitted to firm i via the cost-based Leontief inverse matrix $\tilde{\Psi}$. Recall that $\tilde{\Psi}_{i,r}$ is higher when firm i buys more, directly and indirectly, from firms in region r and when these supplier firms are labor intensive and therefore observe greater decreases in production costs due to immigration.

The first downstream exposure measuring substitution between intermediates

$$D1_{it} = \sum_{j=1}^{n} \frac{\lambda_j}{\lambda_i \mu_j} Cov_{\tilde{\Omega}^{(j)}} \left(\sum_{r=1}^{R} \tilde{\Psi}_{(r)} \delta_{rt}, \Psi_{(i)} \right) \tag{6}$$

summarizes how much firm i's customers (measured by the ith column of $\tilde{\Psi}$) observe cost declines from immigration shock δ_{rt} compared to other firms in the economy. This relates to how much firm i's customers gain or lose business depending on whether different intermediate goods are complements or substitutes.

The second downstream immigration shock capturing substitution between labor and intermediates

$$D2_{it} = \sum_{j=1}^{n} \frac{\lambda_j}{\lambda_i} \tilde{\Omega}_{j,l} \left(\delta_{r_j,t} - \sum_{k=1}^{n} \frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,m}} (\sum_{r=1}^{R} \tilde{\Psi}_{k,r} \delta_{rt}) \right) (\Psi_{ji} - I_{ji}) \tag{7}$$

summarizes how much firm i's customers represented by Ψ_{ji} observe *relative* cost declines from their own region's wages, which is measured by $\delta_{r_j,t}$, compared to the immigration shock through their suppliers, which is measured by $\sum_{k=1}^{n} \frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,m}} (\sum_{r=1}^{R} \tilde{\Psi}_{k,r} \delta_{rt})$.[18]

4.2 Identification Strategy

There are two distinct but equally important identification challenges. The treatments depend on both the immigration intensity across regions δ_{rt} and the shape of the input-output matrix Ω and $\tilde{\Omega}$. Both channels can cause bias in the estimation. Immigrants can choose to locate in regions that have positive labor demand shocks, which would create a positive correlation between trade spillovers of unobserved labor demand and observed labor supply shocks. Alternatively, even if immigrants were randomly assigned across regions, firms that rely more on out-of-region intermediate goods would be more trade-exposed than other firms. If these firms are following different economic trajectories, then comparing more/less trade-exposed firms would also provide biased biased estimates.

[18] One key empirical challenge is to generate the upstream and downstream treatment variables U_{it}, $D1_{it}$ and $D2_{it}$. This requires taking the inverse of huge matrices. For example, at baseline, we have around 230,000 unique firms that trade in the domestic market, resulting in the trade matrices holding approximately 53 billion values. Whereas the trade matrices $\tilde{\Omega}$ and Ω are sparse and therefore do not take up too much memory, the Leontief inverse matrices $\tilde{\Psi}$ and Ψ are not sparse. To make progress, we donated a workstation with 512 GB of RAM to the Ministry of Industry and Technology of Turkey, which hosts most of the datasets we use in this study. Appendix Section A provides the details of how we construct these matrices and our treatment variables.

For identification, we need both quasi-experimental variation in when and where immigrants are settled in, and find firms who are on similar economic trajectories but are differentially exposed to immigrants based on differences in trading partners. We achieve the former by finding an instrument and the latter by using synthetic controls. The resulting estimator is called a Synthetic Instrumental Variable (SIV) estimator (Gulek and Vives-i Bastida, 2024). SIV uses synthetic controls to account for unmeasured confounding while still using the weights assigned by the instrument for identification. Before discussing the implementation, we introduce the instrument.

To generate quasi-experimental variation in immigrant location, we rely on a shift-share instrument, where the share is the average inverse travel distance between Turkish regions and Syrian governorates, and the shift is the aggregate number of Syrians in Turkey. Specifically, our main regional instrument is defined as:

$$Z_{r,t} = \underbrace{\sum_{s=1}^{13} \lambda_s \frac{1}{d_{r,s}}}_{\text{Share}} \times \underbrace{\text{Number of Syrians in Turkey in year } t}_{\text{Shift}} \tag{8}$$

where $d_{r,s}$ is the travel distance between Turkish region r and Syrian governorate s, and λ_s is the weight given to Syrian governorate s.[19] Different weights λ have been used in the literature. In practice, weights matter little. We use the weights suggested by Aksu et al. (2022), which take into account two empirical facts: the number of refugees from a Syrian region s increases with population and proximity to Turkey compared to other bordering countries.

$$\lambda_s = \underbrace{\frac{\frac{1}{d_{s,T}}}{\frac{1}{d_{s,T}} + \frac{1}{d_{s,L}} + \frac{1}{d_{s,J}} + \frac{1}{d_{s,I}}}}_{\text{Relative distance to Turkey}} \times \underbrace{\pi_s}_{\text{Pop. share}} \tag{9}$$

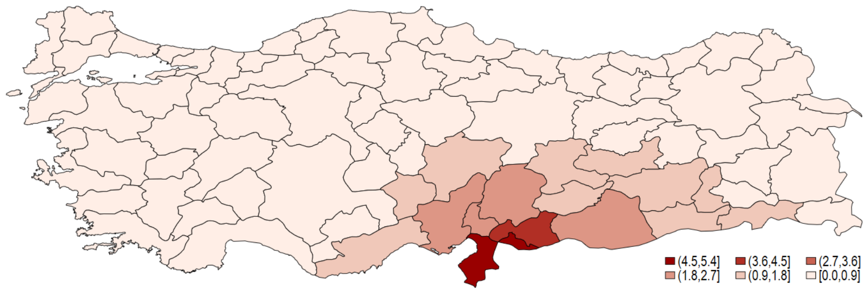
where $d_{s,c} \quad c \in \{T, L, J, I\}$ is the travel distance between Syrian region s to Turkey, Lebanon, Jordan, and Iraq respectively; and π_s is the population share in 2011, which we calculate using the 2011 census undertaken by the Central Bureau of Statistics of Syria.

Figure 3a shows the cross-sectional distribution of the distance share. As expected, the instrument puts more weight on the south-east regions of Turkey that are closer to the north-west part of Syria. This is because in 2011, along the Turkish-Syrian border, more Syrians lived in Aleppo (north-west) than Al-Hasakah (north-east). Figure 3b shows the first-stage estimates from a nonparametric event-study design where we regress the immigration treatment δ_{rt} on the distance-share Z_r interacted with year indicators. Estimates between

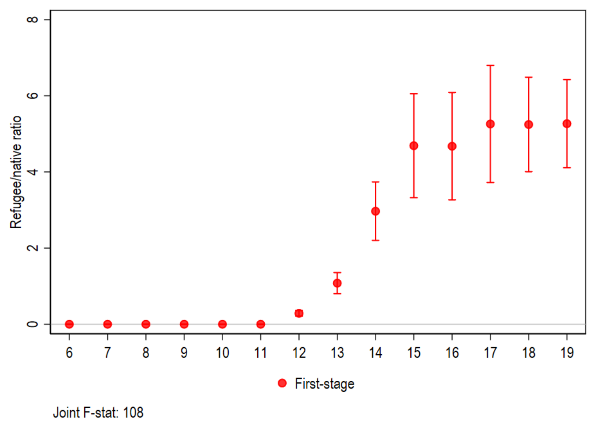
[19] City centers in each region are used to calculate the travel distance. The data is available upon request.

2006–2011 are zero as there are no immigrants in Turkey during those years. In the post-period 2012–2019, distance strongly predicts immigrant location choice in all years. The joint F-stat for the post-period coefficients is 108, which implies that we have a strong instrument.

Figure 3: The Distance instrument



(a) Distance IV heatmap



(b) First-stage

Notes: The heatmap shows the cross-sectional distribution of the distance share Z_r, where the measure is normalized to have unit variance and to start from 0 for the least exposed region. The event-study figure shows the estimates from a nonparametric event-study regression of the first-stage: $\delta_{rt} = \sum_{t' \neq 2011} \beta_{t'} \mathbb{1}\{t' = t\} Z_r + \alpha_r + \alpha_t + \epsilon_{rt}$ where we weight each region by its population in 2011. Standard errors are clustered at zero. 95% confidence intervals are plotted.

As a robustness check, we also show results in the Appendix using an alternative shift-share instrument, where the share is the ratio of Arabic speakers at the province level in the 1965 census. This is similar to the past-settlement instrument of Card (2001), with the main difference being that Arabic-speaking populations were not generated by the past migration of Syrians in Turkey: they are a result of the multi-ethnic population of the Ottoman Empire. Details of this instrument can be found in the Appendix. Our results are robust to using either instrument. We use distance in the main text as it provides a stronger first-stage.

The instruments for trade exposures U, $D1$, and $D2$ are obtained simply by replacing the immigration treatment δ_{rt} with the instrument Z_{rt}. We refer to the instrumented versions of trade exposures by adding an upper-script z: U^z is the instrument for the upstream exposure treatment U. $D1^z$ and $D2^z$ are defined analogously.

4.3 Estimating Equations

IV Design

Given trade exposure treatments U, $D1$, $D2$, and their respective instruments U^z, $D1^z$, and $D2^z$, we define the estimating equations following Propositions 1 and 2 as follows. The estimating equation for the labor share is given by:

$$\begin{aligned} log(LaborShare_{isrt}) &= \beta_1 U_{it} + f_i^L + f_{srt}^L + \theta_1 W_{it}^L + \nu_{it}^L \\ U_{it} &= \gamma_1 Z_{it}^U + g_i^L + g_{srt}^L + \vartheta_1 W_{it}^L + \omega_{it}^L \end{aligned} \tag{10}$$

where $log(Laborshare_{isrt})$ is the natural logarithm of the labor share of firm i in industry s, region r, and at time t, f and g denote the fixed effects in the structural and first-stage equations, respectively, f_i^L and g_i^L denote firm fixed effects, f_{srt}^L and g_{srt}^L denote industry-region-time fixed effects, and ν_{it}^L and ω_{it}^L are the error terms. We include region-industry-time fixed effects to partial out region-industry level shocks such as technology and markup shocks that can be correlated with the treatment.

The non-standard part of this equation is the W_{it}^L term, which is an unobserved confounder that allows for more/less trade-exposed firms to follow different trajectories. For example, in the Appendix we show that firms that are more upstream or downstream exposed to immigrants follow different trends to less-exposed firms in the pre-period, making it unlikely that they would follow similar trends in the post-period absent the immigration shock. For example, firms who buy from immigrant-intensive firms (i.e. firms that are upstream exposed) increase their size, payroll and labor share between 2006–2011 compared to other less-exposed firms even within the same region-industry cells, resulting in significant pre-trends in reduced-form estimates.[20] Consequently, estimating equation 10 using standard IV regressions without controlling for W_{it} would result in inconsistent estimates.

To make progress, we use Synthetic Controls to partial out the unobserved confounders by implementing the SIV procedure. SIV estimator consists of two steps. In the first step, we find synthetic controls for each unit (firm) in the pre-period and generate counterfactual estimates for the outcome, treatments, and instruments. In the second step, as in the

[20] This can be due to various factors, one of which that south-east regions in Turkey observed larger employment gains than the rest of the Turkish economy between 2006–2011 (Gulek, 2024). It is not surprising that these local demand shocks spread through the production network and impact firms in non-host regions.

standard IV estimator, we use these counterfactual estimates to compute the first-stage and reduced-form estimates. We discuss the details of the implementation in Appendix Section C.

In particular, we find the weights by matching the demeaned values of our two target outcomes: the natural logarithms of labor share and sales between 2006-2011.[21] To rely on the variation in treatment between firms in the same region-industry cell for identification, we restrict the donor pool to firms in the same region-industry cell, where industry is defined at the two-digit level. We also add a penalty term á la Abadie and L'hour (2021) to lower over-fitting bias when working with disaggregated data.

The estimating equation for firms' sales is given by:

$$\begin{aligned} log(Sales_{isrt}) &= \beta_2 D1_{it} + \beta_3 D2_{it} + f_i^S + f_{srt}^S + \theta_2 W_{it}^S + \nu_{it}^S \\ D1_{isrt} &= \gamma_2 Z1_{it}^D + \gamma_3 Z2_{it}^D + g_i^S + g_{srt}^S + \vartheta_2 W_{it}^S + \omega_{1,it}^S \\ D2_{isrt} &= \gamma_4 Z1_{it}^D + \gamma_5 Z2_{it}^D + h_i^S + h_{srt}^S + \vartheta_3 W_{it}^S + \omega_{2,it}^S \end{aligned} \tag{11}$$

where the terms are defined analogously to equation 10.

Two points are in order. First, we omit the downstream treatments, $D1$ and $D2$, in equation 10 and the upstream treatment, U, in equation 11, for two reasons. First, these are the correct structural regression equations for identifying the elasticity parameters. Second, as discussed in the identification section, the upstream treatment U is measured with greater precision than the downstream treatments $D1$ and $D2$. Consequently, even though the upstream treatment U does not structurally belong in equation 11, it could absorb the causal effects of the less precisely measured downstream treatments $D1$ and $D2$ if they were estimated jointly. Despite this empirical problem, we show robustness of our main results to jointly estimating the effects of upstream and downstream exposure treatments in the Appendix.

Second, note that our two estimating equations 10 and 11 are linked: both estimate a version of the elasticity of substitution between labor and intermediate goods. Specifically,

$$\beta_1 = -\frac{(1-\sigma_U)}{\epsilon^D} \quad ; \quad \beta_2 = \frac{(1-\sigma_l)}{\epsilon^D} \quad ; \quad \beta_3 = -\frac{(\sigma_l - \sigma_u)}{\epsilon^D}$$

where ϵ_D is the labor demand elasticity with respect to wages, which we calibrate to be -1.27 from Gulek (2024). In the empirical section, we explicitly show that the estimates from our two estimating equations are mutually consistent.

[21] We estimate a common set of weights for both labor share and sales to minimize the noise-to-signal ratio (Sun et al., 2023). Appendix Section C shows that estimating separate weights for labor share and sales results in worse performance of SIV on unmatched outcomes such as payroll and size.

Event-study Design

The primary advantage of the event-study design is that it allows us to visually and flexibly assess the pattern of outcomes the (debiased) share component of the shift-share instruments capture relative to the beginning of the refugee crisis. We define the event-study equations of the SIV estimator for labor share as:

$$\widetilde{log(y_{it}^L)} = \sum_{t' \neq 2011} \beta_{1,t'} \widetilde{U_i^Z} \mathbb{1}\{t = t'\} + f_i^L + f_t^L + \nu_{it}^L \tag{12}$$

and for sales as:

$$\widetilde{log(y_{it}^S)} = \sum_{t' \neq 2011} \left(\beta_{t'}^{D1} \widetilde{D1_i^Z} + \beta_{t'}^{D2} \widetilde{D2_i^Z} \right) \mathbb{1}\{t = t'\} + f_i^S + f_t^S + \nu_{it}^S \tag{13}$$

where the outcomes and the instrument shares are their *debiased* versions from partialing out the region-industry-time fixed effects and the unobserved confounder.

4.4 Threats of Identification

There are a few threats to identification that are worth discussing. First, evidence from 10 is likely to be more credible from the evidence from 11 due to two separate but equally important issues: noise and informality in the sales data of small firms. First, sales information λ comes from balance sheet records. Due to the low audit probability of small firms, balance sheet sales are highly noisy. This noise enters both the outcome, lowering precision, and the downstream exposure definitions, causing attenuation bias.

The second problem due to informality is more nuanced. Gulek (2024) and Bahar et al. (2024) show that informal immigration episodes increase firms' labor informality in host regions. Informal workers are paid in cash, which itself often comes from informal transactions that do not appear in Balance sheet or VAT data. As host region firms' demand for informal workers increases, their demand for informal transactions may increase. Consequently, both their purchases from and sales to non-host regions may disappear from the data in the post-period.

We address these problems in several ways. To address attenuation bias, we define our baseline exposure variables by averaging sales and costs between 2006—-2011 instead of relying on data from any particular year. Averaging across years lowers the noise embedded in the data-generating process and, hence, should lower the bias from noise. To address the potential biases from informal sales, we show evidence separately for large firms (50+ employees in 2010) as informality rates decrease with firm size in Turkey. Third, we also show the effect of downstream exposure on employment and exports, the former because it is less noisy and the incentives to hire workers informally do not change in non-host regions, and the latter because it cannot be as easily hidden compared to domestic sales.

Another concern is that trade exposure treatments capture the direct effect of immigration on host regions. This could happen, for example, if the trade exposures were correlated with immigration intensity within region-industry cells. For example, large firms trade more across regions, and these firms rely less on immigrant labor. To address this problem, we drop from the estimating sample all the firms in regions where immigrants constitute more than 4% of the native population and regions that received a large weight by the instrument. Appendix Figure B.1 shows the areas that are dropped from estimation.

4.5 Reduced Form and 2SLS estimates

Cost Propagation

We begin by estimating the reduced-form effects of upstream exposure on firms' labor demand. Specifically, we estimate equation 12 and plot the results in Figure 4. The outcome variable is the number of employees in Figure 4a and the total payroll in Figure 4b. There are four main takeaways from Figure 4a, which displays the estimated effects of upstream exposure on firms' size.

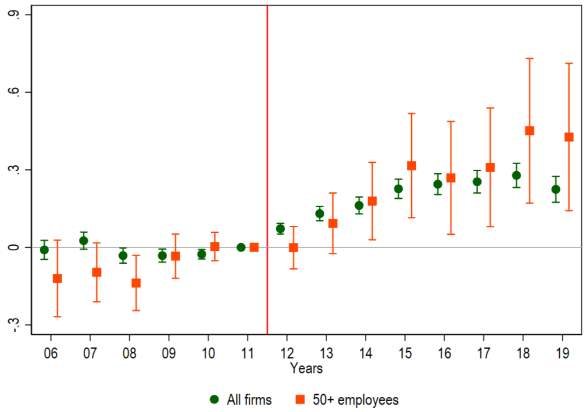
First, we do not see statistically or economically significant pre-trends. This is strong evidence in favor of our identification strategy. Recall that SIV weights were generated to match the trends in labor share and sales, not payroll or firm size. Therefore, the lack of pre-trends in Figure 4a is not mechanical. It shows evidence of a common underlying factor that generates differential trends between more/less exposed firms, and that SIV is able to partial out this unobserved confounder.

Second, upstream exposure significantly increases firms' size. Put differently, firms in non-host regions who directly or indirectly buy from immigrant-intensive firms in host regions hire more workers. Moreover, the magnitudes of the estimated effects increase over time, similar to the first-stage shown in Figure 3b, which improves our confidence that the estimated effects are causal effects and not differential trends.

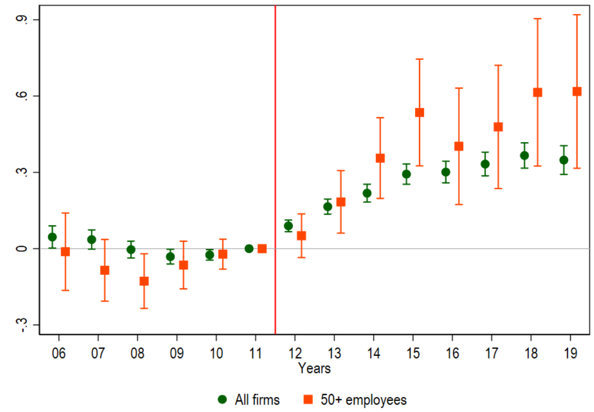
Third, estimates from the sample of only large firms are less precise because of the decreased sample size. This is a trade-off between bias and variance. Large firms are less informal and their data is arguably more credible, but there are fewer of them to obtain precise estimates.[22]

[22]Among manufacturing firms that survive throughout 2006–2019, only 6.5% have 50+ employees at baseline.

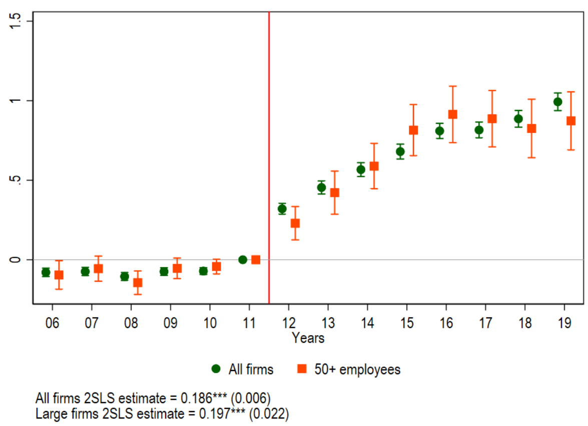
Figure 4: Effect of Upstream Exposure on Firms' Labor Demand



(a) Number of employees



(b) Payroll



(c) Labor share

Notes: The estimates come from the regression equation $\widetilde{y_{it}} = \sum_{t' \neq 2011} \gamma_{1,t'} \widetilde{U_i^Z} \mathbb{1}\{t = t'\} + f_i + f_t + \nu_{it}$, where the outcome variable is the natural logarithm of the number of workers in Panel A, of total payroll in Panel B, and of labor share in Panel C. Both the outcome and the treatment are their debiased versions following the SIV algorithm. In each panel, regression estimates from two separate samples are plotted: one involving firms of all sizes, and one involving only firms with at least 50 employees at baseline. The upstream exposure is given by $U_i^Z = \sum_{r=1}^{R} \tilde{\Psi}_{i,r} Z_r$, where $\tilde{\Psi}$ is the cost-based Leontief inverse matrix, and Z_r is the regional share of the instrument. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

Fourth, despite differences in precision, estimates using all firms and only large firms are economically and statistically similar to each other. This means that upstream exposure increases the size of both small and large firms in similar magnitudes. These effects could be different if, for example, the production technologies of small and large firms were different. The estimates imply that the elasticities of substitution between labor and intermediate goods are similar across both types of firms.

Interpreting the coefficients in this reduced-form design is not straightforward as the treatment is a general equilibrium exposure. Consider two firms, which we denote by i_1 and i_2. Both firms spend half of their costs on labor and the other half on one intermediate good. Firm i_1 buys from firm j_1, and firm i_2 buys from firm j_2. Further suppose that firms j_1 and j_2 also use half of their costs in labor. Let j_1 be two standard deviations more exposed to immigrants through distance than firm j_2. As all firms have a labor share of 1/2, the difference in the upstream exposures of their customers i_1 and i_2 is 1/2 units. The .22 estimate in Figure 4a by year 2019 in Panel A means that firm i_1 increases its size by 11% compared to firm i_2.

Figure 4b shows similar evidence on the effects of upstream exposure on firms' payroll. Coefficient estimates are near zero before the immigration shock. Estimates from the post-period are positive and statistically significant for both small and large firms. Notice that the effects on payroll are slightly larger than the effects on size. As payroll is equal to the number of workers multiplied by the average salary of workers, this evidence shows that upstream exposure weakly increases wages paid to workers.

Figure 4c shows the effects of upstream exposure on firms' labor share. We do not find significant pre-trends in the data between 2006–2011. As labor share was part of the matching step in calculating Synthetic Control weights, the lack of pre-trends shows good pre-treatment fit in the training period, which is an important condition for SIV to function well. Starting from 2012, we document a significant increase in the labor share of upstream-exposed firms. Firms in the non-host regions who directly or indirectly buy from the host regions increase their labor share compared to other similar firms in the same region-industry cells. In Panel C we also report the 2SLS estimate from equation 10 because these estimates map directly to the structural elasticity parameter between labor and intermediates. The 2SLS estimate from the sample of all manufacturing firms is 0.186. This implies that labor and intermediate goods are gross complements, with an elasticity of substitution of $\sigma_U = 0.75$. The estimates from large firms are highly similar: a 2SLS estimate of 0.197, which implies an elasticity of substitution $\sigma_U = 0.76$.

Labor and intermediate goods are gross complements for all two-digit manufacturing industries. Appendix Figure C.11 shows the estimates of the elasticity of substitution between labor and intermediate goods separately for each two-digit Manufacturing industry. Across 24 manufacturing industries, elasticity estimates range from 0.66 to 0.97, with a median of 0.80. Moreover, they are not statistically different than each other. This finding supports our assumption that the structural elasticities are homogenous across industries.

Before introducing the results of downstream exposure on firms' sales, we discuss the robustness of our estimates of upstream exposure effects. There are in general two types of concerns with SC based estimators: under-fitting, which refers to the inability to find a convex combination of donor units that mimics the treated units, and over-fitting, which refers to SC weights matching on the noise and not the signal in the data. The lack of pre-trends in labor share shows the more exposed firms in our data are not outliers: we are able to generate synthetic firms that follow similar trends. Moreover, the lack of pre-trends in specifications where the outcome variable is size and payroll, which are untargeted outcomes while calculating the SC weights, show evidence against over-fitting.

Another concern with our empirical design is that distance may not be a good instrument for immigrants' location choice. Appendix Figure C.7 replicates Figure 4 using the alternative language instrument. We find similar results.

To sum up, we find that upstream exposure increases firms' labor demand, which implies that labor and intermediate goods are complements in production, with an elasticity of substitution of around 0.76. Our results are similar across small and large firms, eliminating concerns related to informality. Quality checks of the SIV estimator show good pre-treatment fit and limited room for bias over-fitting.

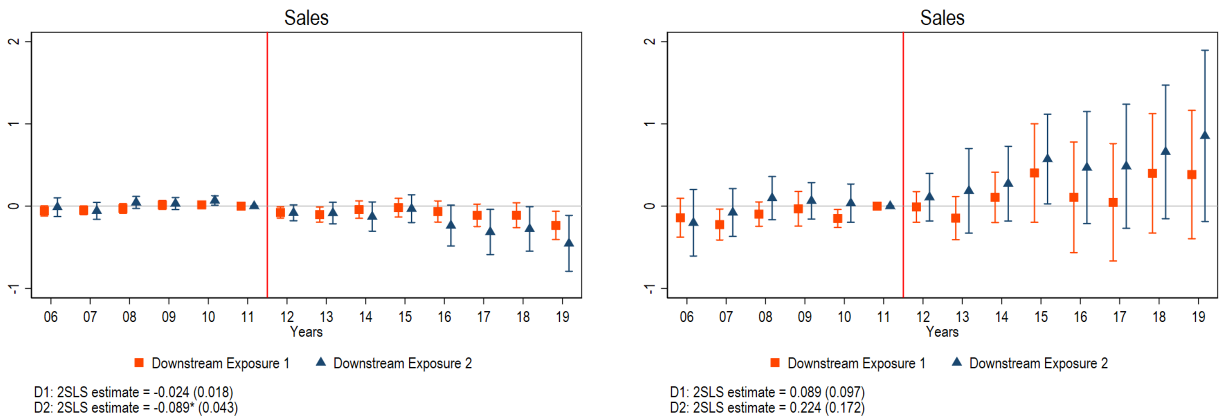
Demand Spillovers

We continue by estimating the reduced-form effects of downstream exposures on firms' sales using SIV. In particular, we estimate equation 13 and plot the results in Figure 5. Figure 5a shows the downstream exposure effects on all manufacturing firms, and Figure 5b shows the downstream exposure effects on large manufacturing firms. Recall that there are two downstream exposure effects: $D1$ is driven by firms' sales exposure to upstream-exposed firms and $D2$ is driven by firms' sales exposure to host regions.

Comparing the estimates of $D1$ and $D2$ effects across small and large firms reveals two important results. First, the effects of $D1$ are small in magnitude and statistically indistinguishable from zero. This is true for both small and large firms. A zero effect of $D1$ exposure means that firms on average do not change the share of expenditures on different intermediate goods in response to the immigration shock. This implies that intermediate goods are neither complements nor substitutes in the aggregate, with an elasticity of substitution of $sigma_L \approx 1$.

Second, comparing the effects of $D2$ between small and large firms shows a dichotomy. Whereas $D2$ exposure lowers firms' sales on average, it increases the sales of large firms. If true, the former would have been a surprising result and a rejection of the model. A negative $D2$ estimate means that labor and intermediates are more substitutable than different intermediates in production. This is inconsistent with both the effects of upstream

Figure 5: Effect of Downstream Exposures on Firms' Sales



(a) All firms

(b) Large Firms

Notes: The estimates come from the reduced-form regression equation $log(\widetilde{Sales_{it}}) = \sum_{t' \neq 2011} \beta_{t'}^{D1} \widetilde{D1_i^Z} \mathbb{1}\{t = t'\} + \beta_{t'}^{D2} \widetilde{D2_i^Z} \mathbb{1}\{t = t'\} + \alpha_i^{Sales} + \alpha_t^{Sales} + \nu_{it}^{Sales}$, where both the outcome and the two treatments are their debiased versions following the SIV algorithm. The downstream exposures are calculated by replacing the immigration treatment δ_{rt} in equations 6 and 7 with the instrument share Z_r. 95% confidence intervals are plotted.

exposure in Figure 4 and prior estimates from the literature (Burstein et al., 2020). In contrast, the evidence from large firms is consistent with our earlier results. The 2SLS estimates among large firms imply an elasticity of substitution between labor and intermediate goods of around 0.83, which is similar to the 0.76 we find from upstream exposure effects.

We perform several robustness checks to ensure that the decrease in sales from $D2$ exposure arises from the aforementioned biases and is not a causal effect. For example, if $D2$ exposure was somehow hurting firms, we would also expect firms to lower their demand for labor. However, Appendix Figure C.8 shows that $D2$-exposed firms do not become smaller. In fact, they increase their labor share in production. They also do not lower exports, which is harder to hide than domestic sales. Overall, the evidence suggests that the estimates in Figure 5a have negative bias due to aforementioned data-related issues, hence we conclude that the evidence from Figure 5b is more credible.

To sum up, we estimate the elasticities of substitution between labor and intermediate goods as $\hat{\sigma_u} = 0.76$, and between different intermediate goods as $\hat{\sigma_l} = 1$. Given that the evidence from large firms across the two structural equations, upstream exposure effect on labor share and downstream exposure effects on sales, are consistent (i.e., the data does not reject the model), we move on to our counterfactual estimates to quantify the total effects of immigration on host and non-host labor markets.

4.6 Quantifying the Aggregate Effects of Immigration

Previous section shows the reduced-form effects of trade exposure on firms in non-host regions. This section uses the model to quantify the aggregate effects of Syrian immigration on the Turkish economy. Recall that Theorem 1 characterizes the general equilibrium effects of an immigration shock on regional wages and firms' prices as a function of the baseline production network and the structural elasticity parameters. We observe the baseline production network in the data and the previous section estimates the structural elasticity parameters. Therefore, solving the system linear equations given in 1 gives us the general equilibrium effects on wages and prices.

One important detail is that wages in the model are defined with respect to nominal GDP, whereas real wages in the real world are usually defined with respect to local prices. Therefore, in this section, we define the change in real wages as $d \ln w_{real} = d \ln w - b * d \ln p$, where b denotes the $R \times N$ matrix of final expenditure shares, and $d \ln p$ is the $N \times 1$ vector of price changes.

We also extend our baseline model to incorporate the empirical fact that Syrian immigrants are predominantly low-skilled compared to the Turkish labor force. The details can be found in the Appendix. While calculating the general equilibrium effects, we assume that all Syrian immigrants were low-skilled.[23] We also assume that low-skill and high-skill workers have an elasticity of substitution of $\sigma_S = 1$.[24]

Figure 6 shows the results. Each circle represents one of the 81 provinces in Turkey. The y-axis denotes the change in the real wages of low-skill natives. The x-axis shows the number of Syrian immigrants per native in 2019. The dashed line is the -45° line. Absent general equilibrium effects, all the blue estimates would be on that dashed line. This figure shows that the general equilibrium effects are highly similar to the partial equilibrium effects. The vector of real wage changes and the immigration shock have a correlation of -0.99, leading to an R-squared of 0.97. Simply put, partial equilibrium effects are a pretty good predictor of the general equilibrium effects.

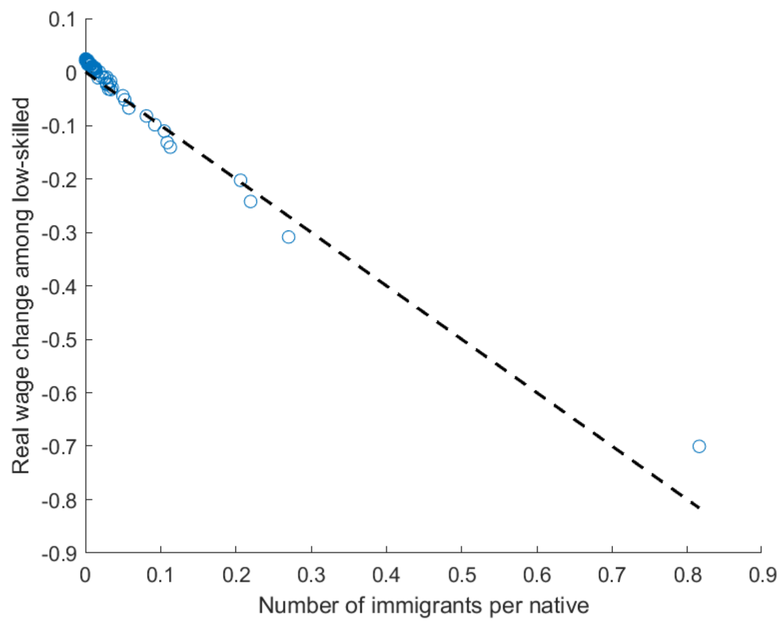
This is a surprising finding given that our reduced form results show robust evidence for spillover effects across regions. What these results tell us is that, whereas some firms in the host regions in Turkey trade across regions, most do not. While our identification strategy is able to isolate these spillover effects, they are negligible in the aggregate.

Is this result generalizable? Are regional spillovers always economically small and negli-

[23] Gulek (2024) shows that Syrian immigrants lower low-skill natives' employment and increase high-skill natives' wages in the exposed industries.

[24] We cannot estimate this elasticity in the data because our employer-employee matched data does not include the education level of workers, and the information on workers' occupation starts in the post-period.

Figure 6: Partial vs General Equilibrum Effects of Syrian Immigration in Turkey



Notes: Provincial distribution of the number of immigrants per native in 2019 is used. The general equilibrium changes in wages and prices are calculated as a solution to the system of linear equations given in the Appendix Section B. Each blue circle denotes a Turkish province. The dashed line is the -45° line.

gible, or is this result specific to the south-east regions in Turkey or to the low-skill intensity of the immigrants? Would there be larger spillover effects if the host regions were different or if immigrants were higher skilled? To answer these questions, we employ counterfactual analysis.

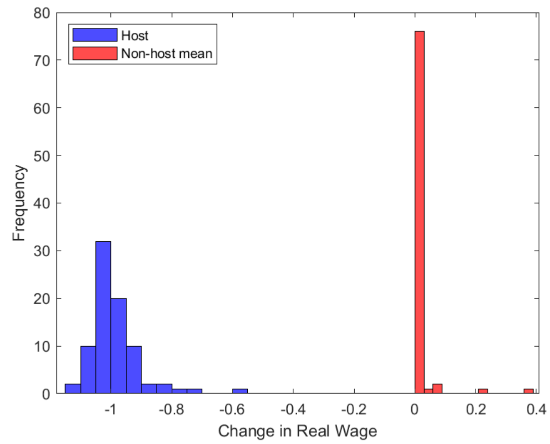
4.7 Counterfactuals

This section uses the model to quantify the effects of immigration on the host and non-host regions through counterfactuals. We investigate the economic significance of the trade spillovers of immigration, how these spillovers depend on host regions' and immigrants' characteristics, and what these spillovers imply about our understanding of the effects of immigration on labor markets.

Counterfactual 1: When do immigrants impact non-host regions in economically meaningful amounts?

In our first counterfactual, we investigate whether immigrants can impact the labor markets in non-host regions in economically meaningful amounts. To answer this question, for each of

Figure 7: Histogram of real wage changes in host and average non-host region



Notes: This figure shows the results from 81 counterfactuals, one for each province in Turkey. Each counterfactual consists of a 1% increase in labor supply in the host province. The "non-host mean" refers to the simple average of real wage changes across the 80 non-host regions. Real wages are calculated by the difference between the change in nominal wages and the change in the regional price index.

the 81 provinces in Turkey, we shock the economy with a 1% increase in labor supply in that region, and calculate the change in real wages in the host region and the average non-host region. The latter is an average of real changes in the other 80 provinces. This gives us 81 different estimates for the real wage effects in the host and non-host regions. Figure 7 shows the histogram of these wage effects. Note that since we assume inelastic labor supply, in an economy where firms only traded within region, a 1% increase in labor supply would have decreased real wages in the host region by 1% and not change the real wages in the non-host regions.

There are two key observations from Figure 7. First, in most counterfactuals, the real wage in the host region decreases by around 1%, and the real wage in the non-host regions does not change on average. In 76 out of 81 host region selections, the real wage in the average non-host region changes by less than 0.02%. This is because, in most regions, firms trade overwhelmingly within region. Consequently, real wage changes in the host region do not lead to economically meaningful price changes in the non-host regions. Put differently, the spillover effects of an immigration shock on real wages are mostly negligible.

Second, for 5/81 host cities, we find economically meaningful spillovers. These are Bursa, Kocaeli, Izmir, Ankara and Istanbul. For example, a 1% increase in labor supply in Istanbul decreases the real wages in Istanbul by 0.56% and increases the real wages in the average non-host region by 0.38%. Similarly, a 1% shock in Ankara decreases real wage in Ankara by 71% and increases the real wage in the average non-host region by 22%.

What explains why these five cities create the largest spillovers to non-host regions? One argument is population: a 1% increase in labor force in Istanbul is seven times as large of immigration shock as a 1% increase in labor force in Gaziantep, one of the major host regions in Turkey. While population certainly plays a role, it is not driver of these results. Kocaeli is less populated than Gaziantep, Sanliurfa and Adana, three major host regions, yet immigration to Kocaeli generates more spillovers than those three cities combined. The GDP of these cities also matters, but does not fully explain the results. Antalya has the same GDP as Kocaeli, yet the spillover effects from Kocaeli are 8 times larger than the spillover effects from Antalya.

What seems to explain these results is the centrality of host regions. In these regions, firms trade substantially more across regions. Based on eigenvalue centrality, Istanbul and Ankara are the most and second most central nodes in the production network. Izmir, Bursa, and Kocaeli follow immediately. This is in line with Acemoglu et al. (2012), who show that shocks affecting sectors that hold more central roles within the production network disproportionately influence overall output.

These results have important implications for the accumulated empirical evidence on the effects of immigration on labor markets and why different studies often find opposing results (Dustmann et al., 2016). The standard way of studying the effects of immigrants on labor markets has been a spatial difference in difference design (DiD), in which regions that receive immigrants are compared to others before and after the shock (Altonji and Card, 1991; Card, 2001). Famously, Card (1990) studied the labor market effects of the Mariel Boatlift on Miami's labor markets by comparing Miami to Atlanta, Houston, Los Angeles and Tampa-St. Petersburg, and found mostly null effects.[25] Identification in this type of DiD analysis relies on the stable unit treatment value assumption (SUTVA), which requires that the labor markets in the "control" (non-host) regions are not impacted by the arrival of immigrants to "treated" (host) regions. Our results show that, when immigrants arrive at the central nodes of the trade network, SUTVA fails to hold.[26] For instance, when Istanbul receives a 1% increase in labor supply, the real wages in otherwise "control" regions increase between 0.30-0.46%, while the real wage in Istanbul goes down by 0.56%. Comparing Istanbul to other regions in Turkey, as is done in DiD analyses, would cause us to massively overestimate the negative impact of immigrants on wages in Istanbul. If the technology parameters were different, for example, if labor and intermediate goods were gross substitutes or if intermediate goods were more substitutable and therefore lead to some firms

[25]Mariel Boatlift increased Miami's labor supply by 7%.

[26]Note that the idea of spatial spillovers of immigration shocks violating SUTVA is not new in the immigration literature. Similar concerns were initially raised by Borjas et al. (1997); Borjas (2003), but the focus was more on natives' ability to move from host to non-host regions in response to immigration.

losing business, then immigrant arrival to Istanbul could have also lowered the real wages in the non-host regions, causing the DiD analysis to underestimate the impact of immigrants on the host region.

Counterfactual 2: Does where immigrants setlle matter for welfare?

Given that spillover effects of immigration vary substantially based on which region receives immigrants, a natural question is whether these differences in spillover effects matter for welfare. What is the optimal allocation of immigrants and refugees across space is an important policy question that several host countries are facing today. Countries including Germany, Sweden, Norway and Finland have policies that direct refugees and asylum seekers to settle in specific regions, usually to prevent overcrowding of big cities. This counterfactual investigates whether there are meaningful welfare gains from overcrowding specific cities that are central in the production network.

To answer this question, we simulate labor supply shocks across the 26 major regions in Turkey equivalent to an arrival of 100,000 immigrants.[27] We calculate the changes in prices throughout the economy for each simulation, and then calculate the aggregate welfare gain in the economy. Specifically, let $d \ln Y_r$ denote the welfare change in region r, which is given by:

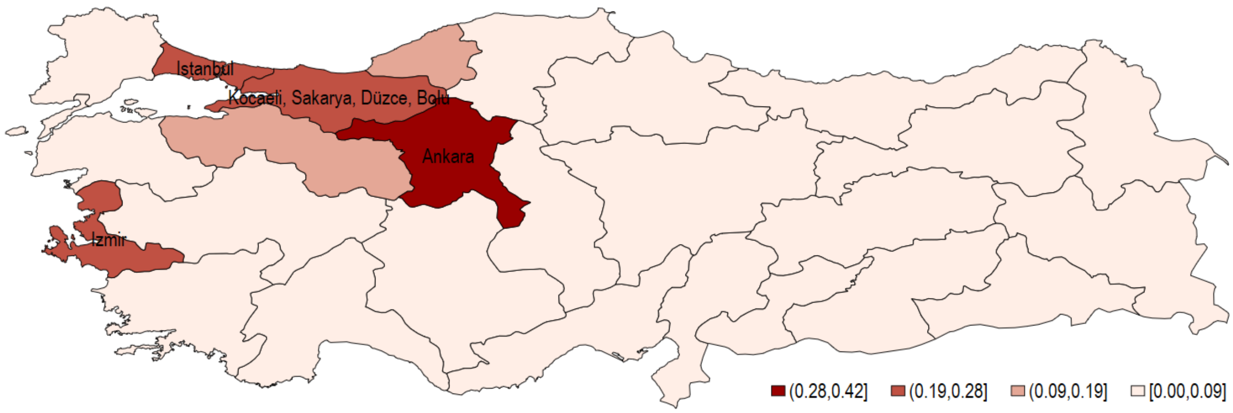
$$d \ln Y_r = d \ln \chi_r - \sum_{i \in N_r} b_i d \ln p_i \tag{14}$$

In words, the representative consumer in region r is better off if its share in total GDP χ_r increases and/or the prices of the goods in its basket decreases. In each simulation, we obtain 26 welfare changes using equation 14. To get at the aggregate welfare change, we take a weighted average across regions where the weight is the population share of each region.

Figure 8 shows the heatmap of the total welfare effects of 100,000 immigrants, a 0.12% increase in total population in Turkey, across different host regions. We see a significant heterogeneity in the total welfare effects of immigration. Whereas immigrants increase total welfare across all 26 trials, it does so little in most regions in Turkey. In 21 out of 26 trials, we document less than 0.09% increase in welfare. In contrast, the welfare effects increase by 0.19-0.42%, up to 21 times larger than the smallest welfare effect of 0.02%, when regions like Izmir, Istanbul, and Ankara receive immigrants. Welfare effects are largest when these cities receive immigrants because they are central nodes in the trade network based on

[27] We use the 26 NUTS-2 regions for this analysis instead of the 81 NUTS-3 regions in Turkey. This is due to the massive heterogeneity in populations across provinces in Turkey (120 thousand in Kilis to 14 million in Istanbul), an immigration shock of the same size creates too large of a difference in the percentage change in local populations.

Figure 8: Heatmap of Total Welfare Effects of Immigration across Host Regions



Notes: This figure shows the results from 26 counterfactuals, one for each NUTS-2 region in Turkey. Each counterfactual consists of an arrival of 100,000 immigrants to the host region. The change in total welfare is calculated by taking a weighted average of the change in regional welfare, where the weights are the share of the population living in that region.

Eigenvector centrality. Firms in these regions buy from and sell to firms in various other regions. Consequently, more regions benefit from the cost reductions, which results in a larger increase in total welfare.

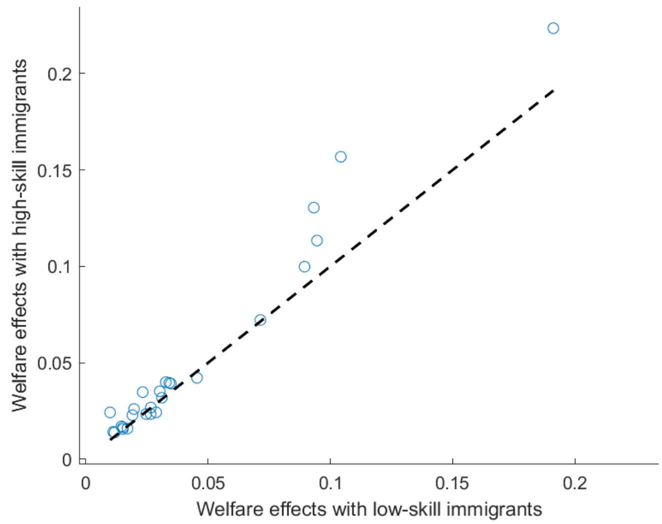
The main takeaway from this counterfactual is that the welfare implications of immigration depend largely on which regions receive immigrants. In the present setting, a social planner that wants to maximize the total welfare in the economy would prefer immigrants to settle at the central nodes of the trade network, which are Izmir, Istanbul, Ankara and the province group consisting of Kocaeli, Sakarya, Duzce, and Bolu.

Counterfactual 3: Does the skill composition of immigrants impact the spillover effects of immigration?

Immigrants and natives can differ in skill levels. For example, Syrian immigrants are less educated than the Turkish native labor force and work in less skill-intensive industries such as Textile, Construction, and Agriculture (Crescent and Programme, 2019). If low-skill and high-skill labor are not perfect substitutes, then low-skill and high-skill immigrants lower production costs in different types of industries. Depending on how much these industries vary in their trade relations with other regions, the spillover effects of immigration can also vary by the skill content of immigrants.

To quantify how much the spillover effects of immigration change based on the skill intensity of immigrants, we run 26 counterfactuals, once for each NUTS-2 region in Turkey, in which we treat the host region with first 100,000 low-skill and then 100,000 high-skill

Figure 9: Comparison of welfare effects across low-skill and high-skill immigration



Notes: This figure shows the results from 52 counterfactuals, two for each NUTS-2 region in Turkey. For each region, we calculate the total welfare change when (1) 100,000 low-skill immigrants arrive in the host region and (2) 100,000 high-skill immigrants arrive in the host region. Low-skill is having less than a high school degree, and high-skill is having at least a high school degree.

immigrants and compare the welfare effects between the two trials.

Figure 9 plots the welfare effects of low-skill and high-skill immigration. Each small circle represents one of the 26 NUTS-2 regions. The x-axis corresponds to the welfare effects of low-skill immigration, the y-axis corresponds to the welfare effects of high-skill immigration, and the dashed line is the 45-degree line. Points above the dashed line are the regions where high-skill immigration leads to higher welfare gains in the overall economy.

There are two takeaways from this figure. First, both low-skill and high-skill immigration create negligible welfare effects for most of our trials. This is consistent with our earlier results, which showed that the spillover effects of immigration, which are inherently linked to the total welfare effects, are negligible when the host regions are not the most developed regions of Turkey. In these cases, it does not matter whether the immigrant is low- or high-skilled: the cost reductions from immigration are contained within region, resulting in negligible welfare effects. In contrast, in cases where the welfare gains are high or, equivalently, the host region is a central node in the trade network, high-skill immigration leads to sizable gains in welfare. For example, an arrival of 100,000 low-skill immigrants increases total welfare in Turkey by 0.10%, as opposed to an arrival of 100,000 high-skill immigrants, which increases total welfare by 0.16%.

5 Conclusion

This paper presents a comprehensive analysis of how immigration-induced wage shocks propagate through regional economies via production networks. The theoretical model and empirical evidence together show that immigration can have significant spillover effects on labor demand, particularly when immigrants settle in central nodes of a domestic trade network and/or when immigrants work in skill-intensive industries. This highlights the importance of considering regional trade structures when evaluating the economic effects of immigration.

Our findings challenge traditional approaches to studying immigration's impact, which often ignore interregional spillovers, and offer new insights into why previous studies may have produced conflicting results. Overall, this paper contributes to the broader literature on immigration and labor markets by demonstrating the importance of incorporating production networks into the analysis, offering policy-relevant insights into the management of large immigration episodes.

References

Abadie, Alberto and Jérémy L'hour, "A penalized synthetic control estimator for disaggregated data," *Journal of the American Statistical Association*, 2021, *116* (536), 1817–1834.

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "Microeconomic origins of macroeconomic tail risks," *American Economic Review*, 2017, *107* (1), 54–108.

_ , David Autor, David Dorn, Gordon H Hanson, and Brendan Price, "Import competition and the great US employment sag of the 2000s," *Journal of Labor Economics*, 2016, *34* (S1), S141–S198.

_ , Ufuk Akcigit, and William Kerr, "Networks and the macroeconomy: An empirical exploration," *Nber macroeconomics annual*, 2016, *30* (1), 273–335.

_ , Vasco M Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "The network origins of aggregate fluctuations," *Econometrica*, 2012, *80* (5), 1977–2016.

Akgündüz, Yusuf, Abdurrahman Aydemir, Seyit Cilasun, and Murat G Kïrdar, "Propagation of Immigration Shocks through Firm-to-Firm Trade Networks," IZA Discussion Paper, IZA 2024.

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.

Altonji, Joseph G and David Card, "The effects of immigration on the labor market outcomes of less-skilled natives," in "Immigration, trade, and the labor market," University of Chicago Press, 1991, pp. 201–234.

Bahar, Dany, Isabel di Tella, and Ahmet Gulek, "Formal Effects of Informal Labor Supply and Work Permits: Evidence from the Venezuelan Refugees in Colombia," 2024. Available at: https://shorturl.at/JwuGW.

Baqaee, David Rezza and Emmanuel Farhi, "The macroeconomic impact of microeconomic shocks: Beyond Hulten's theorem," *Econometrica*, 2019, *87* (4), 1155–1203.

Barrot, Jean-Noël and Julien Sauvagnat, "Input specificity and the propagation of idiosyncratic shocks in production networks," *The Quarterly Journal of Economics*, 2016, *131* (3), 1543–1592.

Boehm, Christoph E, Aaron Flaaen, and Nitya Pandalai-Nayar, "Input linkages and the transmission of shocks: Firm-level evidence from the 2011 Tōhoku earthquake," *Review of Economics and Statistics*, 2019, *101* (1), 60–75.

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

_ , "The wage impact of the Marielitos: A reappraisal," *ILR Review*, 2017, *70* (5), 1077–1110.

_ **and Joan Monras**, "The labour market consequences of refugee supply shocks," *Economic Policy*, 2017, *32* (91), 361–413.

_ **, Richard B Freeman, Lawrence F Katz, John DiNardo, and John M Abowd**, "How much do immigration and trade affect labor market outcomes?," *Brookings papers on economic activity*, 1997, *1997* (1), 1–90.

Brinatti, Agostina, "Third-Country Effects of US Immigration Policy," *Available at SSRN 4892498*, 2024.

Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel, "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence From the United States," *Econometrica*, 2020, *88* (3), 1071–1112.

Caliendo, Lorenzo, Luca David Opromolla, Fernando Parro, and Alessandro Sforza, "Goods and factor market integration: A quantitative assessment of the EU enlargement," *Journal of Political Economy*, 2021, *129* (12), 3491–3545.

Card, David, "The impact of the Mariel boatlift on the Miami labor market," *ILR Review*, 1990, *43* (2), 245–257.

_ , "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

Carvalho, Vasco M, "From micro to macro via production networks," *Journal of Economic Perspectives*, 2014, *28* (4), 23–48.

_ **and Alireza Tahbaz-Salehi**, "Production networks: A primer," *Annual Review of Economics*, 2019, *11* (1), 635–663.

— , **Makoto Nirei, Yukiko U Saito, and Alireza Tahbaz-Salehi**, "Supply chain disruptions: Evidence from the great east japan earthquake," *The Quarterly Journal of Economics*, 2021, *136* (2), 1255–1321.

Crescent, Turkish Red and World Food Programme, "Refugees In Turkey: Livelihoods Survey Findings," 2019.

Dustmann, Christian and Albrecht Glitz, "How do industries and firms respond to changes in local labor supply?," *Journal of Labor Economics*, 2015, *33* (3), 711–750.

— , **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.

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

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *110* (8), 2586–2624.

Gulek, Ahmet, "Formal Effects of Informal Labor: Evidence from the Syrian Refugees in Turkey," 2024. Available at SSRN: https://ssrn.com/abstract=4264865.

— **and Jaume Vives i Bastida**, "Synthetic IV estimation in panels," 2024. Available at: https://economics.mit.edu/sites/default/files/inline-files/Synthetic_Wald%20%283%29.pdf.

Hanson, Gordon H, "The economic consequences of the international migration of labor," *Annu. Rev. Econ.*, 2009, *1* (1), 179–208.

Hunt, Jennifer, "The impact of the 1962 repatriates from Algeria on the French labor market," *ILR Review*, 1992, *45* (3), 556–572.

Lewis, Ethan and Giovanni Peri, "Immigration and the Economy of Cities and Regions," in "Handbook of regional and urban economics," Vol. 5, Elsevier, 2015, pp. 625–685.

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

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

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.

Sun, Liyang, Eli Ben-Michael, and Avi Feller, "Using multiple outcomes to improve the synthetic control method," *arXiv preprint arXiv:2311.16260*, 2023.

A Proofs

Before showing the proofs, we introduce some notation. The trade matrix Ω is of size $(N+R)\times(N+R)$, where the last R rows are zeros. We decompose this matrix as follows.

$$\Omega = \left(\begin{array}{c|c} \Omega^p & \Omega^f \\ \hline 0 & 0 \end{array}\right)$$

where Ω^p denotes the first $N\times N$ portion.

Similarly, the Leontief inverse is defined as

$$\Psi = \left(\begin{array}{c|c} \Psi^p & \Psi^p\Omega^f \\ \hline 0 & I \end{array}\right)$$

where $\Psi^p = (I-\Omega^p)^{-1})$

For ease of notation, we use r only to refer to regions. For example, $\Psi_{i,r}$ refers to ith row and $(N+r)$th column, while $\Psi_{i,j}$ refers to ith row and jth column.

Proof of Proposition 1. The labor share in production of firm i is given by

$$\tilde{\Omega}_{i,L} = \frac{(1-\eta_i)^{\sigma_u} w_r^{1-\sigma_u}}{(1-\eta_i)^{\sigma_u} w_r^{1-\sigma_u} + \eta_i^{\sigma_u} p_{m,i}^{1-\sigma_u}}$$

where $p_{m,i}$ is the price of the CES aggregate intermediate good of firm i. Taking the natural logarithm and differentiating, we get:

$$\begin{aligned} d\ln\tilde{\Omega}_{i,L} &= (1-\sigma_u)d\ln w_r - (1-\sigma_u)\left(\tilde{\Omega}_{i,L} d\ln w_r + \tilde{\Omega}_{i,m} d\ln p_{m,i}\right) \\ &= (1-\sigma_u)(1-\tilde{\Omega}_{i,L})d\ln w_r - (1-\sigma_u)\tilde{\Omega}_{i,m} d\ln p_{m,i}. \end{aligned}$$

Using CES attributes, we can write $d\ln p_{m,i}$ as:

$$\frac{1}{1-\sigma_L}\frac{\sum_{j=1}^n \alpha_{ij}^{\sigma_L}(1-\sigma_L)p_j^{-\sigma_L}dp_j}{\sum_{j=1}^n \alpha_{ij}^{\sigma_l} p_j^{1-\sigma_l}}$$

note that

$$\frac{\alpha_{ij}^{\sigma_l} p_j^{-\sigma_l}}{\sum_{k=1}^n \alpha_{ik}^{\sigma_L} p_k^{-\sigma_L}} = \tilde{\Omega}_{i,j}/(1-\tilde{\Omega}_{i,L})$$

Putting this back into the previous equation, we get:

$$
\begin{aligned}
d\ln\tilde{\Omega}_{i,L} &= (1-\sigma_u)(1-\tilde{\Omega}_{i,L})d\ln w_r - (1-\sigma_u)\sum_{j=1}^{n}\tilde{\Omega}_{i,j}d\ln p_j \\
&= (1-\sigma_u)(1-\tilde{\Omega}_{i,L})\left(d\ln w_r - \sum_{j=1}^{n}\frac{\tilde{\Omega}_{i,j}}{\tilde{\Omega}_{i,L}}d\ln p_j\right)
\end{aligned}
\tag{15}
$$

■

Proof of Proposition 3. Prices are given by $p_i = \frac{\mu_i C_i(p,w,\overline{y}=1)}{A_i}$. Keeping markups and technology constant, $d\ln p_i = d\ln C_i$.

Using Shephard's Lemma, we can show the change in costs as:

$$
\begin{aligned}
d\ln C_i &= d\ln\left(\sum_{j=1}^{n}p_j x_{ij} + w_{r_i}L_i\right) \\
&= \sum_{j=1}^{n}\tilde{\Omega}_{i,j}d\ln p_j + \tilde{\Omega}_{i,L}d\ln w_{r_i}
\end{aligned}
$$

Writing this in vector form, we get:

$$
\begin{aligned}
dlnp &= \tilde{\Omega}d\ln p + \tilde{\Omega}_{,L}.*d\ln w \\
&= \tilde{\Psi}^p(\tilde{\Omega}_{,L}.*d\ln w)
\end{aligned}
$$

which implies

$$
dlnp_i = \sum_{j=1}^{n}\tilde{\Psi}^p_{i,j}\tilde{\Omega}_{j,L}d\ln w_{r_j}
$$

■

Proof of Proposition 2. From accounting identity

$$
\lambda = b'\Psi \leftrightarrow \lambda_i = \sum_{j=1}^{n}b_j\Psi_{ji} = \sum_{j=1}^{n}b_j\Psi_{ji} = \sum_{j=1}^{n}\overline{b_{r_j}}\chi_{r_j}\Psi_{ji}
$$

$$
d\lambda_i = \sum_j \overline{b_{r_j}}d\chi_{r_j}\Psi_{ji} + \sum_j \overline{b_{r_j}}\chi_{r_j}d\Psi_{ji}
\tag{16}
$$

Focusing on the first part of equation 16, we can write $d\chi_{r_j}$ as:

$$\chi_r = sum_{i \in r} \pi_i + w_r L_r$$

which gives

$$d \ln \chi_r = \sum_{i \in r} \frac{\pi_i}{\chi_r} d \ln \lambda_i + w_r L_r d \ln L_r + w_r L_r d \ln w_r$$

Focusing on the second part of equation 16 and using matrix calculus, we can show:

$$d\Psi = \Psi d\Omega \Psi$$

so, we need to get $d\Omega$. First, using CES algebra, we can write

$$\tilde{\Omega}_{i,j} = \frac{1}{A_i} \eta_i^{\sigma_u} \alpha_{ij}^{\sigma_l} p_j^{1-\sigma_l} \overline{p_{m,i}}^{\sigma_l - \sigma_u} \overline{p_{y,i}}^{\sigma_u - 1}$$

Taking the natural logarithm and totally differentiating gives:

$$d \ln \tilde{\Omega}_{i,j} = (1 - \sigma_l) d \ln p_j + (\sigma_l - \sigma_u) d \ln \overline{p_{m,i}} + (\sigma_u - 1) d \ln \overline{p_{y,i}}$$

where $\overline{p_{m,i}}$ is the unit price of intermediate goods for firm i and $\overline{p_{y,i}}$ is the unit price of production for firm i. Rewriting these two terms as functions of changes in wages and intermediate good prices gives

$$\begin{aligned} d \ln \Omega_{i,j} =& (1 - \sigma_l) d \ln p_j + (\sigma_u - 1) \tilde{\Omega}_{i,L} d \ln w_r \\ &+ (\sigma_l - 1 + (1 - \sigma_u)(1 - \tilde{\Omega}_{i,m})) \frac{1}{\tilde{\Omega}_{i,m}} \sum_{k=1}^{n} \tilde{\Omega}_{i,k} d \ln p_k \end{aligned}$$

Collecting terms, one can show

$$d \ln \tilde{\Omega}_{i,j} = (1 - \sigma_l) \left(d \ln p_j - \sum_{k=1}^{n+R} \tilde{\Omega}_{i,k} d \ln p_k \right) + (\sigma_l - \sigma_u) \tilde{\Omega}_{i,L} \left(\frac{1}{\tilde{\Omega}_{i,m}} (\sum_{k=1}^{n} \tilde{\Omega}_{i,k} d \ln p_k) - d \ln w_{r_i} \right)$$

Using $d \ln \Omega_{i,j} = d \ln \tilde{\Omega}_{i,j}$ when markups are constant, and using the covariance term, we get:

$$d\Omega_{i,j} = \frac{1 - \sigma_l}{\mu_i} Cov_{\tilde{\Omega}^{(i)}}(d \ln p, I_{(j)}) + \frac{\sigma_l - \sigma_u}{\mu_i} \tilde{\Omega}_{i,j} \tilde{\Omega}_{i,L} \left(\frac{1}{\tilde{\Omega}_{i,m}} (\sum_{k=1}^{n} \tilde{\Omega}_{i,k} d \ln p_k) - d \ln w_{r_i} \right)$$

From proposition 3, we know

$$d\ln p_i = \sum_{j=1}^{n} \tilde{\Psi}^p_{ij}\tilde{\Omega}_{j,L} d\ln w_{r_j}$$

More succinctly, we can write it as:

$$d\ln p = \sum_{r=1}^{R} \tilde{\Psi}_{(r)} d\ln w_r$$

replacing price changes $d\ln p$ in the equation for $d\Omega_{i,j}$, we get:

$$d\Omega_{i,j} = \frac{1-\sigma_l}{\mu_i} Cov_{\tilde{\Omega}^{(i)}}\left(\sum_g \tilde{\Psi}_{(g)} d\ln w_g, I_{(j)}\right) + \frac{\sigma_l - \sigma_u}{\mu_i}\frac{\tilde{\Omega}_{i,L}}{\tilde{\Omega}_{i,m}}\tilde{\Omega}_{i,j}\left(\sum_{k=1}^{n}\tilde{\Psi}^p_{ik}\tilde{\Omega}_{k,L} d\ln w_{r_k} - d\ln w_{r_i}\right)$$

Using $d\Psi = \Psi d\Omega \Psi$, we get:

$$\begin{aligned} d\Psi_{o,s} =& \sum_{j=1} \frac{\Psi_{o,j}}{\mu_j}(1-\sigma_l) Cov_{\tilde{\Omega}^{(j)}}\left(\sum_g \tilde{\Psi}_{(g)} d\ln w_g, \sum_i I_{(i)}\Psi_{is}\right) \\ &+ \sum_{i=1}^{n} \Psi_{0,i}\frac{\sigma_l - \sigma_u}{\mu_i}\frac{\tilde{\Omega}_{i,L}}{\tilde{\Omega}_{i,m}}\left(\sum_{k=1}^{n}\tilde{\Psi}^p_{ik}\tilde{\Omega}_{kl} d\ln w_{r_k} - d\ln w_{r_i}\right)\sum_{j=1}^{n}\tilde{\Omega}_{i,j}\Psi_{j,s} \end{aligned}$$

Using $d\lambda_i = \sum_j \overline{b_{r_j}} d\chi_r \Psi_{ji} + \sum_j \overline{b_{r_j}}\chi_{r_j} d\Psi_j$ and combining terms, we get:

$$\begin{aligned} d\ln\lambda_i =& (1-\sigma_l)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\frac{1}{\mu_j} Cov_{\tilde{\Omega}^{(j)}}\left(\sum_g \tilde{\Psi}_{(g)} d\ln w_g, \Psi_{(i)}\right) \\ &+ (\sigma_l - \sigma_u)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\frac{\tilde{\Omega}_{j,l}}{\tilde{\Omega}_{j,m}}\left(\sum_{k=1}^{n}\tilde{\Psi}^p_{ik}\tilde{\Omega}_{kl} d\ln w_{r_k} - d\ln w_{r_i}\right)(\Psi_{ji} - I_{ji}) \\ &+ \frac{1}{\lambda_i}\sum_j\sum_r \overline{b_{rj}}\Psi_{ji}\chi_r d\ln\chi_r \end{aligned}$$

where $d\ln\chi$ is given by:

$$d\ln\chi_r = \sum_{i\in r}\frac{\pi_i}{\chi_r} d\ln\lambda_i + w_r L_r d\ln L_r + w_r L_r d\ln w_r$$

This completes the proves of propositions 1, 2 and 3. Theorem 1 is proven directly by these propositions.

$\blacksquare$

B Model with skill heterogeneity

B.1 Setup

The economy consists of N firms indexed by i, R regions indexed by r, where each region is endowed with ℓ_r low-skill and h_r high-skill labor. Each firm operates in one region: r_i denotes the region of firm i. Firms use intermediate goods and local labor in production, and sell their output as both an intermediate good to other producers in all regions and as a final good to local consumers.

Producers

Firm i chooses labor ℓ_i, h_i, and intermediate goods $\{x_{i,j}\}_{j=1}^{n}$ to minimize costs subject to a constant returns nested-CES technology

$$
\begin{aligned}
&\min_{\{x_{ij}\}_{j=1}^{n}, L_i} \sum_{j=1}^{n} p_j x_{ij} + w_{r_i,\ell}\ell_i + w_{r_i,h}h_i \quad \text{subject to} \\
&A_i(\eta_i m_i^{\frac{\sigma_u - 1}{\sigma_i}} + (1-\eta_i)L_i^{\frac{\sigma_u-1}{\sigma_u}})^{\frac{\sigma_u}{\sigma_u-1}} \geq y_i \\
&m_i = \left(\sum_{j=1}^{n} \alpha_{ij} x_{ij}^{\frac{\sigma_m-1}{\sigma_m}}\right)^{\frac{\sigma_m}{\sigma_m-1}} \\
&L_i = \left(\alpha_{i\ell}\ell_i^{\frac{\sigma_L-1}{\sigma_L}} + (1-\alpha_{i\ell}h_i^{\frac{\sigma_L-1}{\sigma_L}}\right)^{\frac{\sigma_L}{\sigma_L-1}}
\end{aligned}
$$

where A_i is a Hicks-neutral productivity shifter, y_i is total output, p_j is the price of good j, ℓ_i and h_i are the low-skill and high-skill labor used by firm i, $w_{r,l}$ and $w_{r,h}$ are the low-skill and high-skill wages in region r, m_i is the intermediate good used by the firm, which itself is a CES bundle of goods from different firms. x_{ij} denotes how much firm i uses firm j's goods in production, where firm j can be in any region. We assume common elasticities of substitution within nests: σ_u denotes the elasticity of substitution between labor and intermediate goods, unlike the text, σ_m is the elasticity of substitution between different intermediate goods, and $\sigma_L$is the elasticity of substitution across labor. Constant returns to technology requires $\sum_j \alpha_{i,j} = 1$. Let C_i denote the unit cost of firm i. We assume that firms have constant and exogenous markup μ_i, and therefore set price $p_i = \mu_i C_i$.

Final Demand

All final goods consumption as well as the ownership of firms is local. We assume a representative consumer in each region r, who optimizes her Cobb-Douglas utility subject to budget constraint that equates her spending on final goods with her labor income plus (regional) firm profits.

$$\max_{\{c_{r,i}\}} \Pi_{i\in r} c_{r,i}^{\beta_i} \quad s.t. \quad \sum_{i\in r} p_i x_{0,i} = w_{r,l} l_r + w_{r,h} h_r + \sum_{i\in r} \pi_i$$

where $c_{r,i}$ is how much the representative agent r consumes firm i's goods, and $\sum_{i\in r} \beta_i = 1$.

Labor Supply

Both types of labor are inelasticly supplied in each region, are immobile across regions and perfectly mobile across firms in a region.

General Equilibrium

Given exogenous productivities A_i and markups μ_i, equilibrium is a set of prices p_i, low-skill wages $w_{r,l}$ and high-skill wages $w_{r,h}$, intermediate good choices $x_{i,j}$, labor input choices l_i, outputs y_i, and final demands $c_{r,i}$, such that each producer minimizes its costs subject to technology constraints and charges the relevant markup on its marginal cost; consumers maximize their utility subject to their budget constraint, and the markets for all goods and labor clear.

B.2 Input-Output definitions

We use the same notation as in the baseline model. The only difference worth noting is that the trade matrix Ω is of dimension $(N + 2R) \times (N + 2R)$, where the first N columns and rows belong to firms, rows $N+1$ to $N+R$ belong to low-skill labor, and $N+R+1$ to $N+2R$ belong to high-skill labor.

Effects of a Labor Supply Shock on labor income

Note that the labor income in region r is the sum of labor payments by all firms in that region.

$$\lambda_r = l_r w_{rl} + h_r w_{rh} = \sum_{i\in r} \lambda_i \Omega_{i,L}$$

which gives

$$d\ln\lambda_r = \frac{1}{\lambda_r}\left(\ell_r w_{rl}(d\ln l_r + d\ln w_{rl}) + h_r w_{rh}(d\ln h_r + d\ln w_{rh})\right)$$
$$d\ln\lambda_r = s^w_{LS}(d\ln\ell_r + d\ln w_{rl}) + s^w_{hs}(d\ln h_r + d\ln w_{rh})$$

where s denotes shares of low-skill and high-skill labor expenses.

Proposition 4 characterizes the change in labor share as a function of changes in wages and prices

Proposition 4. *In response to an immigration-induced wage shock, the following equation describes the change in the labor share of production costs*

$$d\ln\tilde{\Omega}_{i,L} = (1-\sigma_u)\left(\frac{1-\tilde{\Omega}_{i,L}}{\tilde{\Omega}_{i,L}}(\tilde{\Omega}_{il}d\ln w_{rl} + \tilde{\Omega}_{ih}d\ln w_{rh}) - \sum_{j=1}^{n}\tilde{\Omega}_{ij}d\ln p_j\right) \tag{17}$$

Proof follows the same steps as in the baseline model.

Let $\overline{w_j} = \frac{1}{\tilde{\Omega}_{j,L}}\left(\tilde{\Omega}_{j,l}d\ln w_{r_j,l} + \tilde{\Omega}_{j,h}d\ln w_{r_j,h}\right)$ is the change in the average wage for firm j. This allows us to write the change in firms' sales using the notation from the baseline model.

Proposition 5. *In response to an immigration-induced wage shock, the following equation describes the change in the Domar weights / sales share of firms*

$$\begin{aligned} d\ln\lambda_i = &\sum_{j=1}^{n}(1-\sigma_M)\frac{\lambda_j}{\lambda_i\mu_j}Cov_{\tilde{\Omega}^{(j)}}\left(d\ln p, \Psi_{(i)}\right) \\ &+ (\sigma_U - \sigma_M)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\tilde{\Omega}_{j,l}\left(d\ln\overline{w_{r_j}} - \sum_{k=1}^{n}\frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}}d\ln p_k\right)(\Psi_{ji} - I_{ji}) \\ &+ \frac{1}{\lambda_i}\sum_j\sum_r\overline{b_{rj}}\Psi_{ji}\chi_r d\ln chi_r \end{aligned} \tag{18}$$

where I is the identity matrix, and $d\ln\chi_r = \left(\sum_{i\in r}\frac{\pi_i}{\chi_r}d\ln\lambda_i\right) + \frac{\lambda_r}{\chi_r}d\ln\lambda_r$ is the change in regional income.

Proof follows the same steps as in the baseline model.

The following characterizes the change in prices.

Proposition 6. *In response to an immigration-induced wage shock, the following equation describes the change in prices charged by firms*

$$d\ln p_i = \sum_{j=1}^{n}\tilde{\Psi}^p_{i,j}(\tilde{\Omega}_{j,l}w_{r_j,l} + \tilde{\Omega}_{jh}w_{r_j,h}) = \sum_{f=1}^{F}\tilde{\Psi}_{i,f}d\ln w_f \tag{19}$$

where f denotes factors, which are the low and high-skill labor in regions.

Proof follows the same steps as in the baseline model.

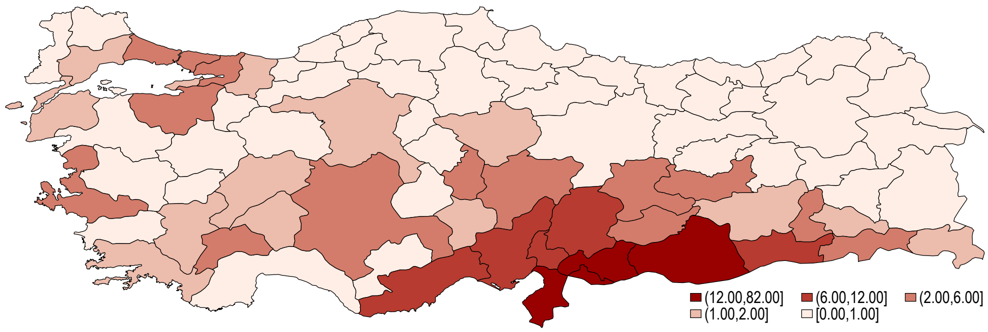
With these propositions at hand, we can fully characterize the effect of an immigration shock on wages and prices.

Theorem 2. *The following linear system fully describes the change in equilibrium prices and quantities in response to an immigration shock consisting of $d\ln l_r$ change in low-skill labor and $d\ln h_r$ change in high-skill labor.*

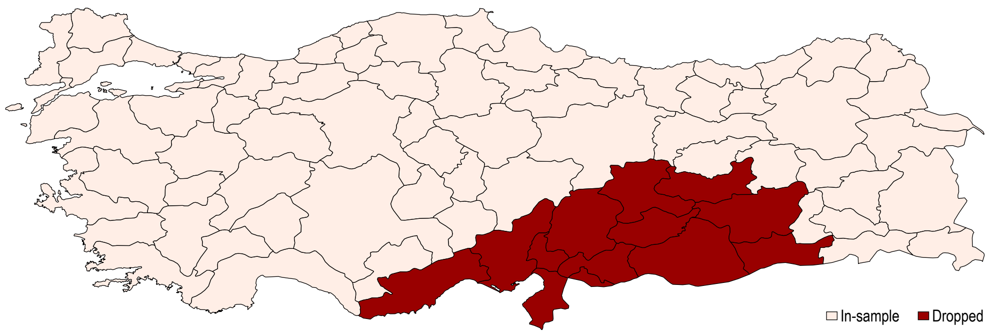
$$
\begin{aligned}
d\ln w_f =& d\ln\lambda_f - d\ln L_f \\
d\ln p_i =& \sum_{f=1}^{F}\tilde{\Psi}_{i,f} d\ln w_f \\
d\ln\lambda_i =& \sum_{j=1}^{n}(1-\sigma_M)\frac{\lambda_j}{\lambda_i\mu_j}Cov_{\tilde{\Omega}^{(j)}}\left(d\ln p, \Psi_{(i)}\right) \\
&+ (\sigma_U-\sigma_M)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\tilde{\Omega}_{j,l}\left(d\ln\overline{w_{r_j}} - \sum_{k=1}^{n}\frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}}d\ln p_k\right)(\Psi_{ji}-I_{ji}) \\
&+ \frac{1}{\lambda_i}\sum_{j}\sum_{r}\overline{b_{rj}}\Psi_{ji}\chi_r d\ln\chi_r \\
d\ln\chi_r =& \left(\sum_{i\in r}\frac{\pi_i}{\chi_r}d\ln\lambda_i\right) + \frac{\lambda_r}{\chi_r}d\ln\lambda_r \\
d\ln\lambda_r =& s^w_{LS}d\ln\lambda_{rl} + s^w_{hs}d\ln\lambda_{rh} \\
d\ln\lambda_f =& \sum_{i=1}^{n}\frac{\lambda_i\Omega_{if}}{\lambda_f}d\ln\lambda_i + \sum_{i=1}^{n}\frac{\lambda_i\Omega_{if}}{\lambda_f}d\ln\Omega_{if} \\
d\ln\tilde{\Omega}_{il} =& d\ln\tilde{\Omega}_{iL} + (1-\sigma_L)\left[d\ln w_{r_i,l} - \frac{1}{\tilde{\Omega}_{iL}}\left(\tilde{\Omega}_{il}d\ln w_{r_i,l} + \tilde{\Omega}_{ih}d\ln w_{r_i,h}\right)\right] \\
d\ln\tilde{\Omega}_{ih} =& d\ln\tilde{\Omega}_{iL} + (1-\sigma_L)\left[d\ln w_{r_i,h} - \frac{1}{\tilde{\Omega}_{iL}}\left(\tilde{\Omega}_{il}d\ln w_{r_i,l} + \tilde{\Omega}_{ih}d\ln w_{r_i,h}\right)\right] \\
d\ln\tilde{\Omega}_{iL} =& (1-\sigma_u)\left[\frac{1-\tilde{\Omega}_{iL}}{\tilde{\Omega}_{iL}}(\tilde{\Omega}_{il}d\ln w_{rl} + \tilde{\Omega}_{ih}d\ln w_{rh}) - \sum_{j=1}^{n}\tilde{\Omega}_{ij}d\ln p_j\right]
\end{aligned}
\tag{20}
$$

Proof follows the same steps as for the baseline model.

Figure B.1: Omitted Regions



(a) Number of refugees per 100 natives in 2019



(b) Regions omitted from the main analysis

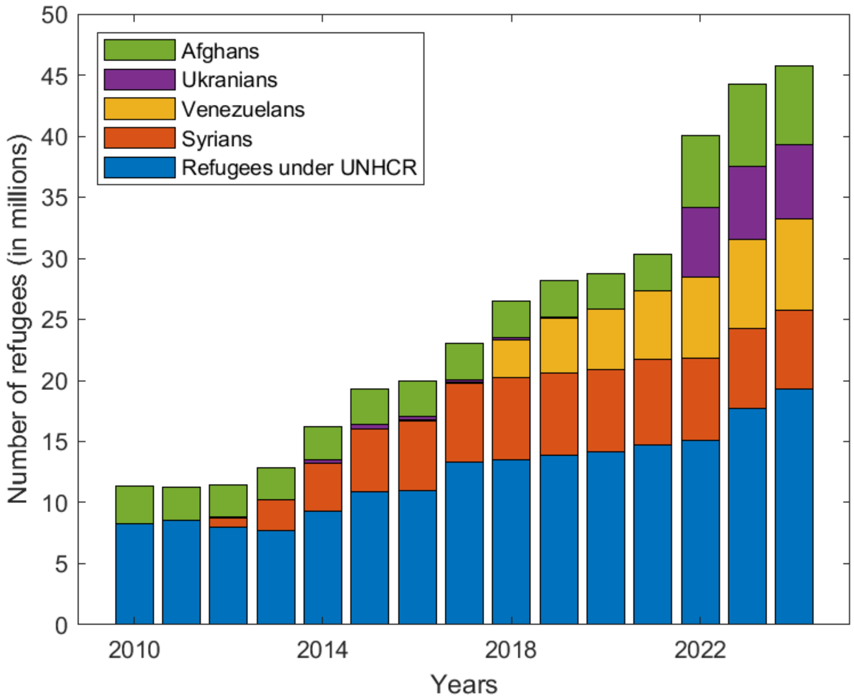
B.3 Data Appendix

Table B.1: Summary Statistics

Number of employees	Wage Bill (in million)	Sales (in million)	Exporter	Labor Share
Panel A: All sizes				
33.11	0.52	7.4	0.27	0.31
(172.49)	(4.84)	(164)	(0.44)	(0.32)
Panel B: More than 20 employees in 2010				
100.46	1.69	28	0.56	0.2
(313.95)	(8.98)	(256)	(0.49)	(21)
Panel C: More than 50 employees in 2010				
217.74	4.00	68	0.71	0.16
(495.77)	(14.51)	(419)	(0.45)	(0.15)

Note: Data is resticted to Manufacturing firms in non-exposed regions that exist throughout 2006–2019. There are 19505 such firms in the sample. 4633 of these firms have more than 20 employees in 2010, and 1112 have more than 50. Standard deviations in parentheses.

Figure B.2: The Evolution of the Number of Refugees Globally



Source: Author's Calculations

C Supporting Evidence

C.1 Comparisons between IV and SIV

The main text emphasizes that more and less exposed firms in the same region-industry cells were on different economic trajectories before the immigration shock. This section shows evidence for these claims.

Specifically, we define the event-study equations of the IV estimator for labor share as:

$$log(y^L_{isrt}) = \sum_{t' \neq 2010} \beta_{1,t'} U^Z_i \mathbb{1}\{t = t'\} + f^L_i + f^L_{srt} + W^L_{it} + \nu^L_{it} \tag{21}$$

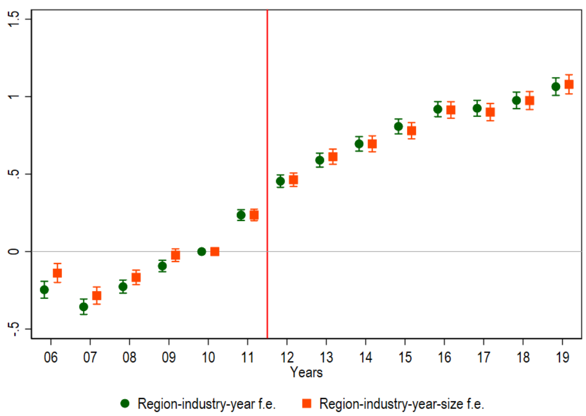
and for sales as:

$$log(y^S_{isrt}) = \sum_{t' \neq 2010} \left(\beta^{D1}_{t'} D1^Z_i + \beta^{D2}_{t'} D2^Z_i\right) \mathbb{1}\{t = t'\} + f^S_i + f^S_{srt} + W^S_{it} + \nu^S_{it} \tag{22}$$

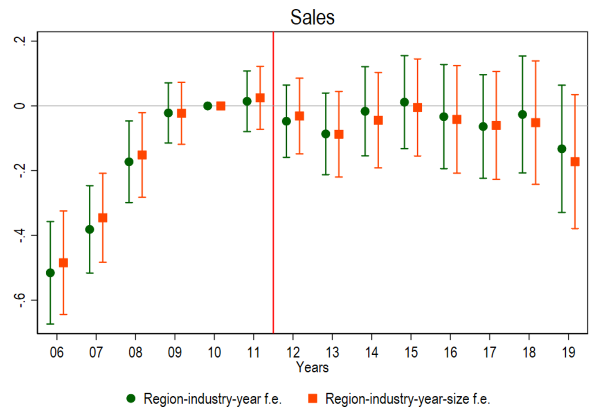
where f_{srt} denotes industry-region-time fixed effects that partial out industry-region level shocks. In robustness checks, we also group firms into quartiles based on their sizes at baseline and control for industry-region-size-time fixed effects. In this specification, we compare firms within the same region-industry cell who have similar number of employees at baseline.

Figure C.3 shows the results. Panel A shows the estimated upstream exposure effects on labor share, and Panels B and C show the estimates downstream exposure effects on sales. Looking at Panel A, one can see that more upstream-exposed firms were following a differential trend from 2007 to 2011 compared to less-exposed firms. The differential increase in labor share from 2007 to 2011 is similar to the increase from 2011 to 2016. This is true even when we control for region-industry-size-time fixed effects. Similarly, Panel B also shows significant pre-trends in the reduced-form with baseline IV. D1-exposed firms' sales grew between 2006–2009 compared to less exposed firms. These differential trends are the reason why we employ Synthetic IV in the main text.

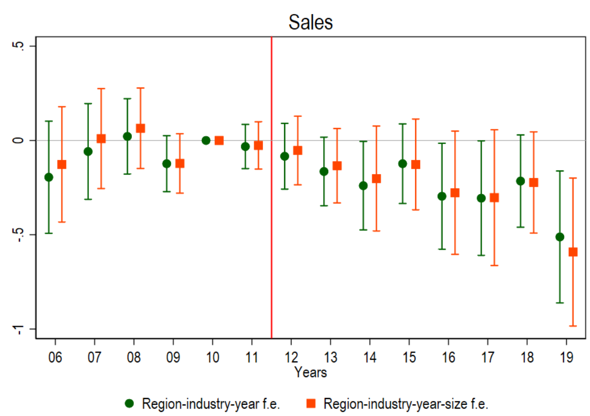
Figure C.3: IV-based Reduced-form Estimates of Upstream and Downstream Exposures on Firms' Labor Demand



(a) Upstream exposure



(b) Downstream exposure 1



(c) Downstream exposure 2

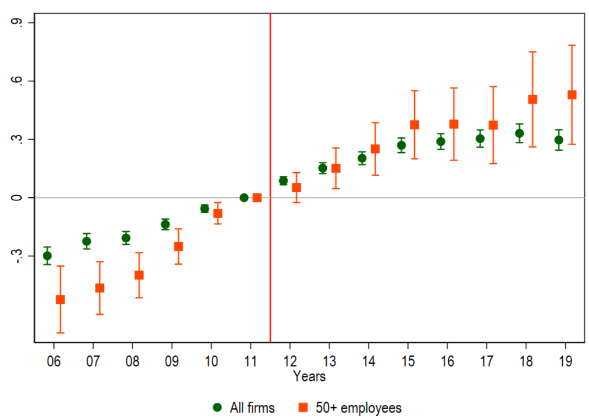
Notes: The estimates in Panel A come from the regression equation $log(y^L_{isrt}) = \sum_{t' \neq 2011} \beta_{1,t'} U^Z_i \mathbb{1}\{t = t'\} + f^L_i + f^L_{srt} + W^L_{it} + \nu^L_{it}$, where the outcome variable is the natural logarithm of the labor share. The estimates in Panels B and C come from the regression $log(y^S_{isrt}) = \sum_{t' \neq 2011} \left(\beta^{D1}_{t'} D1^Z_i + \beta^{D2}_{t'} D2^Z_i\right) \mathbb{1}\{t = t'\} + f^S_i + f^S_{srt} + W^S_{it} + \nu^S_{it}$, where the outcome variable is the natural logarithm of sales. W_{it} denote the region-industry-size-time fixed effects, where size is the quartiles of the number of employees at baseline. Estimates with and without W are plotted. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

C.2 Matching on labor share and sales separately

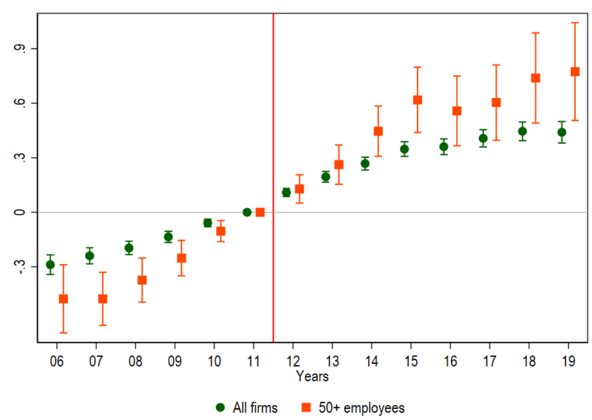
The main text argues that whie applying SIV, calculating the SC weights my matching on the trends in labor share and sales jointly, and therefore having only one set of SC weights for all the outcomes in the it the study, performs better than calculating weights separately for each outcome. The latter strategy suffers from overfitting. Here, we show evidence for our claims.

First, we show the evidence when we match only on labor share. Figure C.4 displays the effects of upstream exposure on the number of employees in Panel A and on payroll in Panel B. We see economically and statistically significant pre-trends in the estimates on firm size and payroll. Among firms that follow similar trends in labor share within the same region-industry cell, those who are more and less upstream-exposed to immigrants follow different trends. This is true for both small and large firms.

Figure C.4: Pre-trends in Upstream Exposure Design when SC weights match only on Labor Share



(a) Number of employees

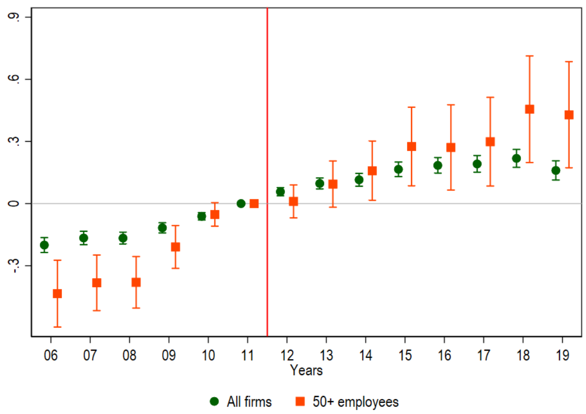


(b) Payroll

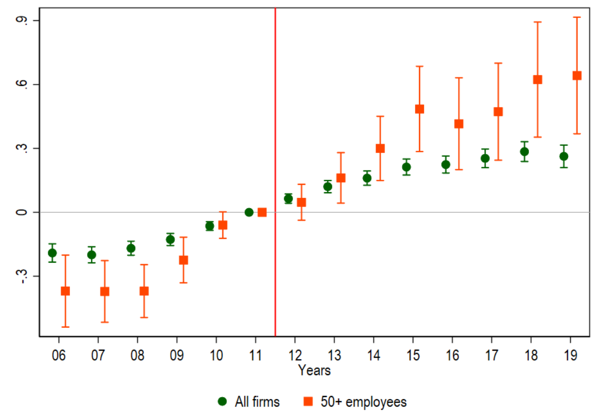
Notes: The estimates come from the regression equation $\widetilde{y_{it}} = \sum_{t' \neq 2011} \gamma_{1,t'} \widetilde{U_i^Z} \mathbb{1}\{t = t'\} + f_i + f_t + \nu_{it}$, where the outcome variable is the natural logarithm of the number of workers in Panel A, of total payroll in Panel B, and of labor share in Panel C. Both the outcome and the treatment are their debiased versions following the SIV algorithm. Unlike the main text, SC weights are calculated by mathing only on the trend in labor shares. In each panel, regression estimates from two separate samples are plotted: one involving firms of all sizes, and one involving only firms with at least 50 employees at baseline. The upstream exposure is given by $U_i^Z = \sum_{r=1}^{R} \tilde{\Psi}_{i,r} Z_r$, where $\tilde{\Psi}$ is the cost-based Leontief inverse matrix, and Z_r is the regional share of the instrument. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

Second, we show the evidence when we match only on the trends in sales. Figure C.5 displays the effects of upstream exposure on the number of employees in Panel A and on payroll in Panel B. We see economically and statistically significant pre-trends in the estimates on firm size and payroll. Among firms that follow similar trends in sales within the same region-industry cell, those who are more and less upstream-exposed to immigrants follow different trends. This is true for both small and large firms.

Figure C.5: Pre-trends in Upstream Exposure Design when SC weights match only on Sales



(a) Number of employees



(b) Payroll

Notes: The estimates come from the regression equation $\widetilde{y_{it}} = \sum_{t' \neq 2011} \gamma_{1,t'} \widetilde{U_i^Z} \mathbb{1}\{t = t'\} + f_i + f_t + \nu_{it}$, where the outcome variable is the natural logarithm of the number of workers in Panel A, of total payroll in Panel B, and of labor share in Panel C. Both the outcome and the treatment are their debiased versions following the SIV algorithm. Unlike the main text, SC weights are calculated by mathing only on the trend in sales. In each panel, regression estimates from two separate samples are plotted: one involving firms of all sizes, and one involving only firms with at least 50 employees at baseline. The upstream exposure is given by $U_i^Z = \sum_{r=1}^{R} \tilde{\Psi}_{i,r} Z_r$, where $\tilde{\Psi}$ is the cost-based Leontief inverse matrix, and Z_r is the regional share of the instrument. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

C.3 Language Instrumental Variable

Immigrants choose where to locate based on local labor market conditions, which implies that our regional immigration treatment δ_{rt} can be correlated with unobserved shocks to labor demand. To address this issue, in the main text we rely on a distance-based shift-share design. A core part of the identification strategy is based on distance shares creating exogenous variation in where immigrants settle. In this section, we show that our main results hold even when we use an alternative instrument for immigrants' location choice.

In this section, we rely on a shift-share instrument, where the share is the ratio of Arabic speakers at the province level in the 1965 census, and the shift is the aggregate number of Syrians in Turkey. This is similar in essence to the past-settlement instrument of Card (2001), with the main difference being that Arabic speaking populations were not generated by the past migration of Syrians in Turkey: they are a result of the multi-ethnic population of the Ottoman Empire. Similar to past-settlement, ethnic similarity is also a strong predictor of where immigrants locate within Turkey.

$$Z_{r,t} = \underbrace{\text{Ratio of Arabic speakers in 1965}}_{\text{Share}} \times \underbrace{\text{Total number of Syrians in Turkey}}_{\text{Shift}} \tag{23}$$

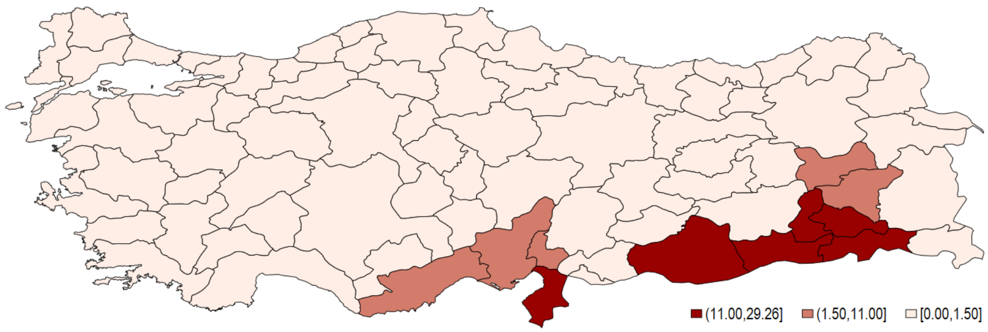
Appendix Figure C.6a shows the cross-sectional distribution of the Arabic speakers in 1965 in Turkey, and Figure C.6b shows the first-stage estimates in an event-study design. Overall, the instrument has a weaker first-stage, of around 10, than the distance instrument, which has a first-stage of around 100.

We define the upstream and downstream instruments the same way as described in the main text. We simply change the shares from distance-shares to language-shares.

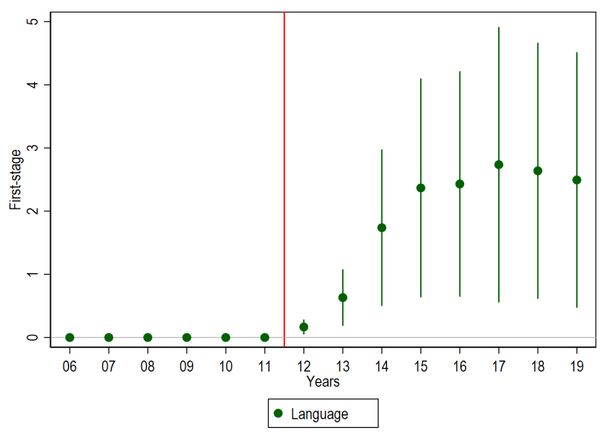
Once we have the language instruments, we estimate the reduced-form of the SIV design. Figure C.7 shows the estimated effects of upstream exposure on firms' size, payroll and labor share. Overall, we find similar results with the main sample of all manufacturing firms. No pre-trends between 2006–2011, and positive increases in firms' labor demand in the post period.

One inconsistency with the main results is for large firms: we cannot statistically distinguish from zero the effect of upstream exposure on employment. The estimates are positive but imprecise. This is because the language instrument has a smaller first-stage than the distance instrument. Using a weaker IV results in less precise estimates.

Figure C.6: Language Instruments

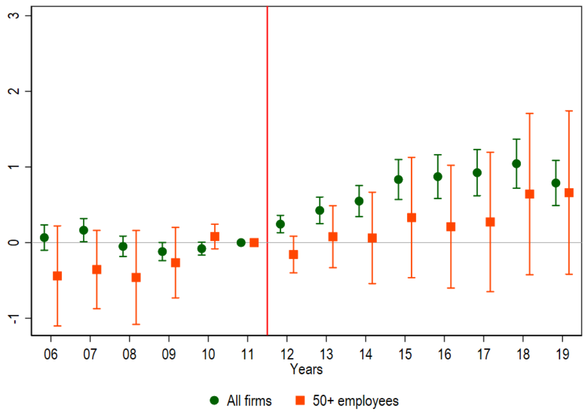


(a) Language Exposure

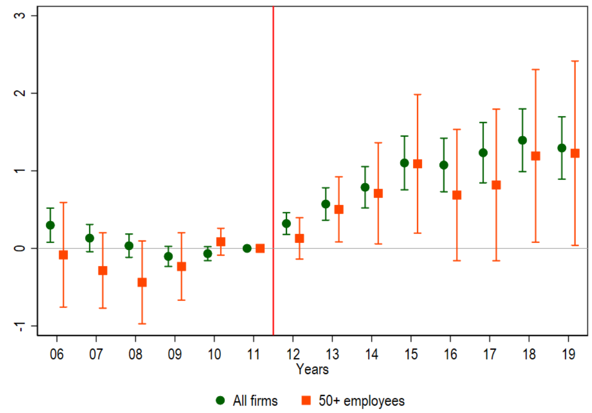


(b) First-stage

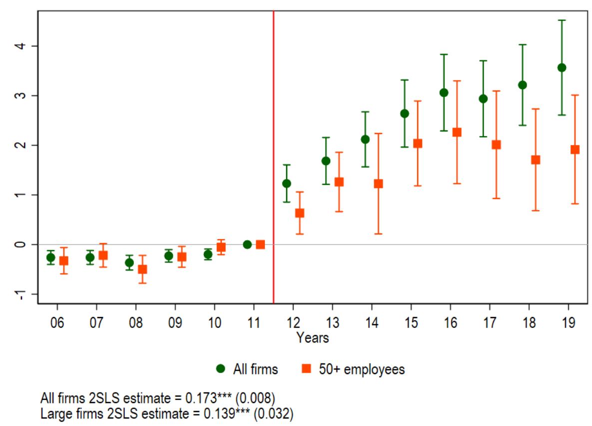
Figure C.7: Effect of Upstream Exposure on Firms' Labor Demand (Language IV)



(a) Number of employees



(b) Payroll



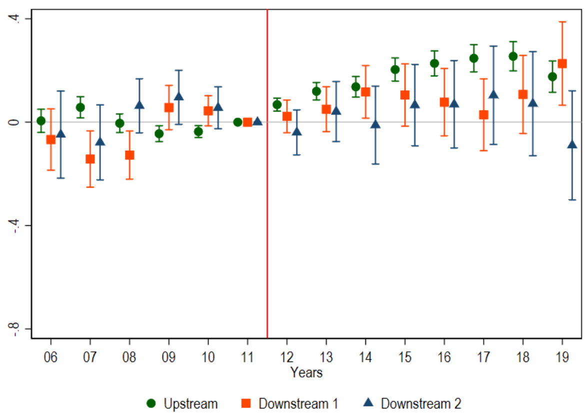
(c) Labor share

Notes: The estimates come from the regression equation $\widetilde{y_{it}} = \sum_{t' \neq 2011} \gamma_{1,t'} \widetilde{U_i^Z} \mathbb{1}\{t = t'\} + f_i + f_t + \nu_{it}$, where the outcome variable is the natural logarithm of the number of workers in Panel A, of total payroll in Panel B, and of labor share in Panel C. Both the outcome and the treatment are their debiased versions following the SIV algorithm. The instruments are based on the Language IV instead of the Distance IV. In each panel, regression estimates from two separate samples are plotted: one involving firms of all sizes, and one involving only firms with at least 50 employees at baseline. The upstream exposure is given by $U_i^Z = \sum_{r=1}^{R} \tilde{\Psi}_{i,r} Z_r$, where $\tilde{\Psi}$ is the cost-based Leontief inverse matrix, and Z_r is the regional share of the instrument. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

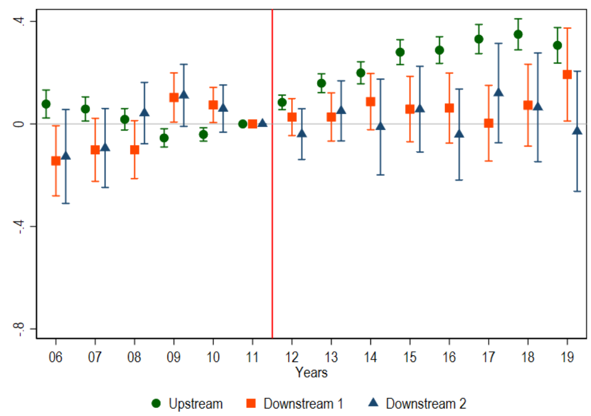
C.4 Joint Estimation of Upstream and Downstream Trade Exposures

In the main text we show that for small firms, downstream exposure appears as if it is decreasing firms' sales. We claim that this is not a true causal effect: downstream exposure does not reduce firms' sales, it reduces their observed sales. To show evidence for this claim, we show that downstream exposure does not lower firms' labor demand. We estimate the effect of upstream and downstream exposures on firms' labor demand (number of employees, payroll and labor share) and plot the results in Figure C.8. As we can see, $D2$ exposure does not lower firms' labor demand. If anything, it leads to an increase in the labor share.

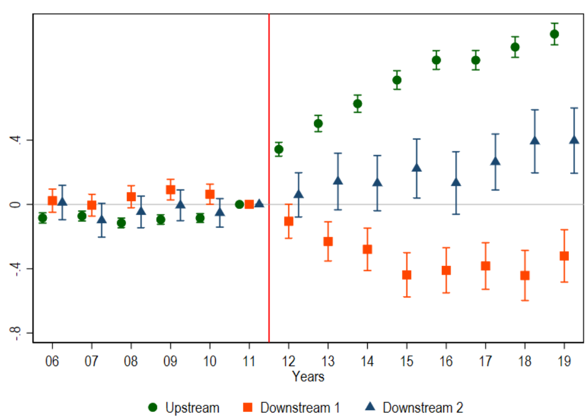
Figure C.8: Effect of Upstream Exposure on Firms' Labor Demand



(a) Number of employees



(b) Payroll



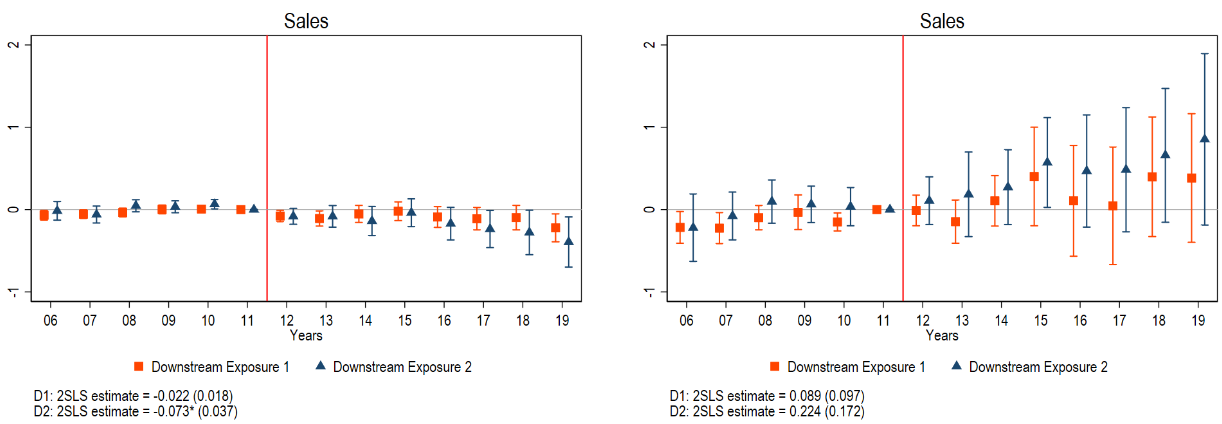
(c) Labor share

Notes: The estimates come from the regression equation $\widetilde{y_{it}} = \sum_{t' \neq 2011} \left(\gamma_{1,t'} \widetilde{U_i^Z} + \gamma_{2,t'} \widetilde{D1_i^Z} \gamma_{3,t'} \widetilde{D2_i^Z} \right) \mathbb{1}\{t = t'\} + f_i + f_t + \nu_{it}$, where the outcome variable is the natural logarithm of the number of workers in Panel A, of total payroll in Panel B, and of labor share in Panel C. Both the outcome and the treatment are their debiased versions following the SIV algorithm. In each panel, regression estimates from two separate samples are plotted: one involving firms of all sizes, and one involving only firms with at least 50 employees at baseline. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

C.5 Noise correction via data cleaning

In the main text, we state that a potential criticism to the results in Figure 5b is that the effects of $D2$-exposure are not statistically significant. We claimed that this is due to the noise in the sales data even among large firms. Here, we show that further restricting the sample to those with less noisy sales data at baseline reveal marginally significant effects, which are shown in the Appendix Figure C.9.

Figure C.9: Effect of Downstream Exposures on Firms' Sales (data restricted to firms with at least 5000 TRY sales at baseline)



(a) All firms

(b) Large Firms

Notes: The estimates come from the reduced-form regression equation $\widetilde{log(Sales_{it})} = \sum_{t' \neq 2011} \beta_{t'}^{D1} \widetilde{D1_i^Z} \mathbb{1}\{t = t'\} + \beta_{t'}^{D2} \widetilde{D2_i^Z} \mathbb{1}\{t = t'\} + \alpha_i^{Sales} + \alpha_t^{Sales} + \nu_{it}^{Sales}$, where both the outcome and the two treatments are their debiased versions following the SIV algorithm. The downstream exposures are calculated by replacing the immigration treatment δ_{rt} in equations 6 and 7 with the instrument share Z_r. 95% confidence intervals are plotted.

C.6 Native Migration Responses

In the text, we argue that Turkish natives do not move in meaningful amounts due to the Syrian immigration shock. To show this, we estimate the following event-study and IV designs at the province level. The event-study design is given by

$$y_{rt} = \sum_{t'=1}^{T} \beta_{t'} Z_r \mathbb{1}\{t = t'\} + f_r + f_t + \epsilon_{rt} \tag{24}$$

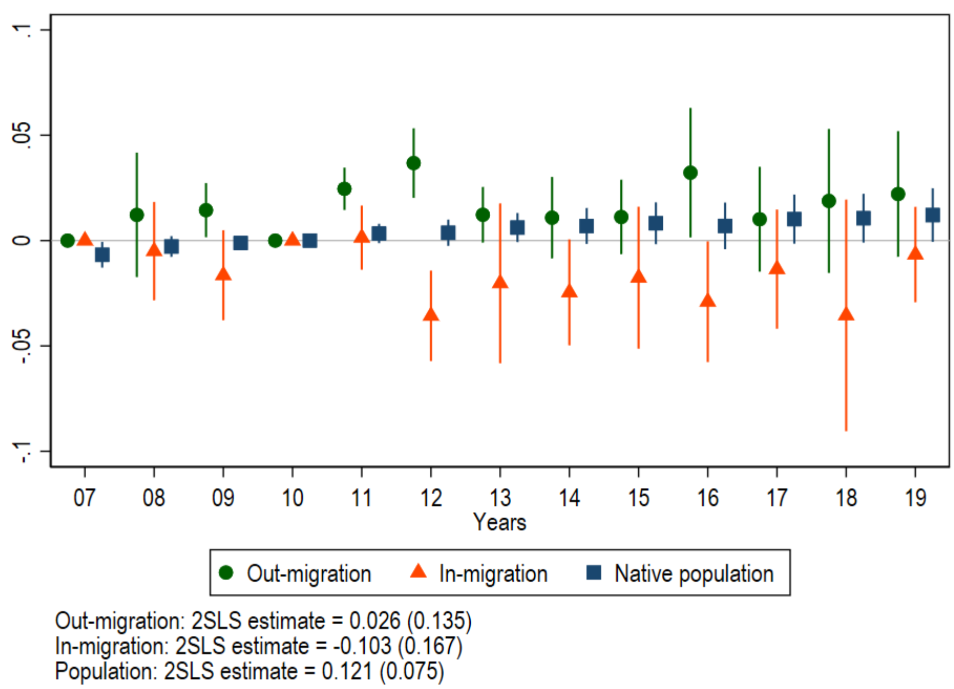
where f_r and f_t are region and time fixed effects. Similarly, the IV design is given by

$$\begin{aligned} y_{rt} &= \beta D_{rt} + f_r + f_t + \epsilon_{rt} \\ D_{rt} &= \gamma Z_{rt} + g_r + g_t + \eta_{rt} \end{aligned}$$

where D is the immigration treatment, Z is the instrument, and f_r, f_t, g_r, g_t are region and time fixed effects. We estimate these designs for three separate outcomes: (the natural logarithms of) in-migration, out-migration and population. Figure C.10 shows the results. Overall, there is a mild decrease in in-migration rates and a mild increase in out-migration rates that were prevalent more in the earlier years of the Syrian civil war (2011 and 2012) before Syrian immigrants started to arrive in masses. In later years, the estimates are not statistically different than zero, and their magnitudes are low. For example, one standard deviation increase in the instrument, which leads to around 9% increase in immigrant/native ratio by 2018, leads to 4% decrease in in-migration rates by 2018. In-migration rates constitute less than 3% of the local population in the host regions. Even if this was a statistically significant effect, magnitude-wise we would conclude that a 1% increase in immigrant/native ratio decreases the native population by around 0.01%. Similar arguments can be made for the out-migration effects.

The effects on in-migration and out-migration rates are so low that the native population continues its upwards trajectory in south-east regions. Due to higher birthrates, south-east regions in Turkey have a higher increase in local population before the Syrian crisis began. This upward trajectory continues in the post period despite the arrival of Syrian immigrants. We conclude that native labor movements across regions does not play a significant role in the dissemination of the immigration shock across regions.

Figure C.10: Native migration responses to Syrian immigration



Notes:

C.7 Industry Heterogeneity

In the main text, we argue that structural elasticity estimates are common across industries. Here, we provide the empirical evidence. To estimate the elasticity of substitution between labor and intermediate goods across different industries, we simply estimate equation 10 separately for each two-digit manufacturing industry. This provides 24 separate SIV estimates of, using which we calculate the structural elasticities. The additional empirical challenge here is that by dividing the data into smaller subgroups, we lower the sample size and therefore the statistical power for each parameter estimate. Just by pure randomness, one could find heterogenous treatment effects when the null is homogenous effect. To account for this, I employ Empirical Bayes Shrinkage.

Let β_j be the elasticity of substitution estimate for industry j. Let $\hat{\beta}_j$ be an estimate of β_j. Assume that the identification strategy is correct, hence $\hat{\beta}_j$'s are consistent estimators of unknown β_j's:

$$\hat{\beta}_j|\beta_j \sim N(\beta_j, s_j^2)$$

Let F denote the distribution of industry-specific EoS occupation-specific child penalties. Suppose F is a normal distribution and independent of s_j's. This gives the following hierarchical model:

$$\begin{aligned}\hat{\beta}_j|\beta_j, s_j &\sim N(\beta_j, s_j^2)\\ \beta_j|s_j &\sim N(\mu_\beta, \sigma_\beta^2)\end{aligned}$$

In this normal/normal model, the posterior mean and variance for β_j given $\hat{\beta}_j$ is given by

$$\begin{aligned}\beta_j^* &\equiv E[\beta_j|\hat{\beta}_j] = \left(\frac{\sigma_\beta^2}{\sigma_\beta^2 + s_j^2}\right)\hat{\beta}_j + \left(\frac{s_j^2}{\sigma_\beta^2 + s_j^2}\right)\mu_\beta\\ s_j^{2*} &\equiv E[s_j^2|\hat{s}_j^2] = \frac{s_j^2\sigma_\beta^2}{s_j^2 + \sigma_\beta^2}\end{aligned}$$

We use the following estimators for the hyperparameters $\mu_\beta, \sigma_\beta^2$.

$$\begin{aligned}\hat{\mu}_\beta &= \frac{1}{J}\sum_{j=1}^{J}\hat{\beta}_j\\ \hat{\sigma}_\theta^2 &= \frac{1}{J}\sum_{j=1}^{J}\left[(\hat{\beta}_j - \hat{\mu}_\beta)^2 - s_j^2\right]\end{aligned}$$

Replacing the unknown parameters by their estimates, we obtain the Empirical Bayes posterior mean and variance:

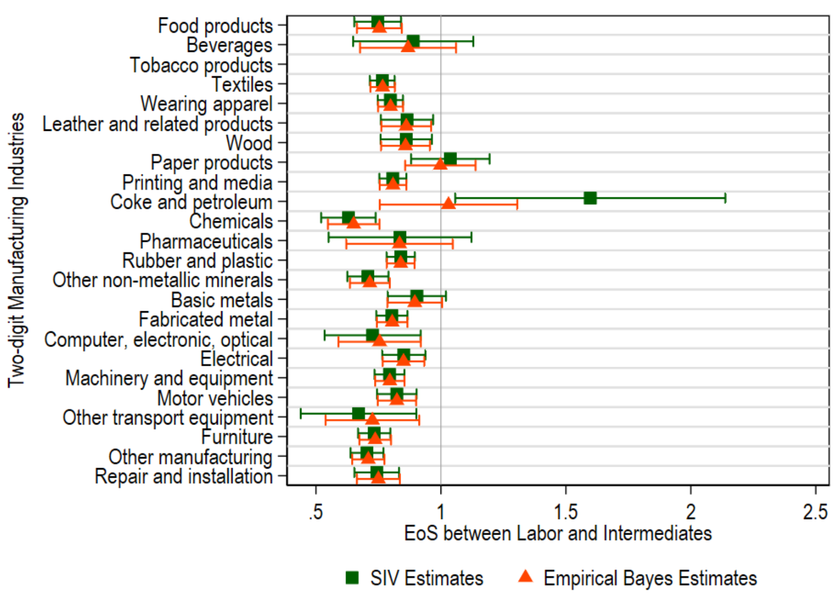
$$\hat{\beta}_j^* = \left(\frac{\hat{\sigma}_\beta^2}{\hat{\sigma}_\beta^2 + s_j^2}\right)\hat{\beta}_j + \left(\frac{s_j^2}{\hat{\sigma}_\beta^2 + s_j^2}\right)\hat{\mu}_\beta$$

$$\hat{s}_j^{2*} = \frac{\hat{s}_j^2\hat{\sigma}_\beta^2}{\hat{s}_j^2 + \hat{\sigma}_\beta^2}$$

Figure C.11 plots both the SIV and the EB posterior estimates of the structural elasticity of substitution between labor and intermediate goods. Panel A shows the estimates using the distance instrument, and Panel B shows the estimates using the language instrument. The evidence is highly similar across the two measures. Notice that SIV and EB estimates are similar except for Coke and Petroleum. This is because the SIV estimates are precise compared to the observed variation in point estimates across industries. Therefore, EB updating assigns most of the weight to the data and less of the weight to the prior. The EB estimates using the distance instrument ranges from 0.66 Chemicals to 0.97 in Coke and petroleum.

Figure C.11: Heterogeneity of EoS between labor and intermediates across Manufucturing industries

(a) Distance IV



(b) Language IV