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

Ahmet Gulek*

October 26, 2025

Most recent draft here.

Abstract

I study how local immigration shocks impact labor markets and firms across the economy through production networks. Using Turkey's Syrian refugee crisis and firm-level trade network data, I show that firms buying from host regions demand more labor, while those selling to host regions increase sales. These spillovers depend critically on network centrality: a 1% labor supply increase in Istanbul decreases local real wages by 0.56% while increasing non-host wages by 0.38%. For non-central regions, identical shocks reduce local wages by 1% with negligible spillovers. Network position thus determines whether immigration only lowers local wages or also generates economy-wide gains.

JEL Classification: F16, F22, J15, J23, J61

Keywords: Immigration, trade, production network

*Postdoctoral Scholar, Immigration Policy Lab, Stanford University, 450 Jane Stanford Way, Stanford, CA 94305–2004, USA (e-mail: agulek@stanford.edu). Tishara Garg was a co-author on an earlier version of the paper, and her contribution is gratefully acknowledged. I am specially indebted to Daron Acemoglu, Josh Angrist, and Amy Finkelstein. I am also thankful to Ran Abramitzky, David Autor, Arnaud Costinot, Christian Dustmann, Jens Hainmüller, Adam Solomon, Edward Wiles, Emma Wiles, participants in the Barcelona Summer Forum Migration, RFBerlin-CEPR Annual Symposium in Labour Economics, The Sixth World Labor Conference hosted by SOLE, Econometric Society World Meeting in Seoul, 8th Lindau Meeting in Economic Sciences, 18th AFD-World Bank Migration and Development Conference at Paris School of Economics, the Stanford IPL Seminar, the MIT Labor/PF Seminar, Labor Lunch and Trade Lunch. I acknowledge support from the George and Obie Shultz Fund at MIT, and from Institute of Humane Studies Grant no. IHS019109.

1 Introduction

How does immigration affect native workers? Three decades of research following Card (1990) have compared regions receiving immigrants to those that do not, yet the literature still lacks consensus on even the sign of these effects.[1] In this paper, I argue that a critical piece of the puzzle has been overlooked: the impact of immigration on non-host regions through production networks. When immigrants reduce production costs in host regions, these effects propagate to firms throughout the economy via input-output linkages—potentially biasing traditional estimates and fundamentally altering our understanding of the aggregate effects of immigration on labor markets. Using the Syrian refugee crisis in Turkey, I present a tractable theory formalizing these mechanisms, empirical evidence testing their existence, and counterfactual exercises quantifying when network spillovers matter for wages and welfare across the economy.

There are three key economic mechanisms by which a local immigration shock propagates through the supply network to impact labor demand throughout the economy. First, immigrants reduce wages, which in turn lowers the prices charged by firms in the host region. This reduction in prices propagates forward to firms that 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. Immigrants' effects also propagate backwards in two distinct ways, which I label as "downstream exposure" effects. If intermediate goods are gross substitutes, then firms whose production costs fall more sharply gain market share. Consequently, they demand more from their suppliers, who observe increases in sales. Furthermore, when intermediates are more substitutable with other intermediates than with labor, immigrant-intensive firms increase their demand for intermediates, which creates a positive demand spillover for their suppliers. Together, these three economic forces shape the labor market effects of immigrants in both host and non-host regions.

My model captures these mechanisms through two key features. First, firms combine local labor with intermediate inputs using constant elasticity of substitution (CES) production technology, where intermediate inputs themselves are CES aggregates of goods from firms across all regions. Second, firms set prices using exogenous markups, which ensures that changes in production costs, whether from labor or intermediate inputs, are passed through to prices. The general equilibrium effects of immigration on labor demand across regions are governed by two key parameters: the elasticity of substitution (EoS) between labor and intermediates, and the EoS across different intermediates. Combined with the structure of the input-output network, these elasticities are sufficient to determine how immigration-induced wage changes in host regions affect labor demand throughout the economy.

The Syrian refugee crisis in Turkey provides an ideal setting for estimating these elasticities. Between 2011 and 2019, Turkey received 3.6 million Syrian refugees, generating labor supply shocks

[1] See Borjas (2017); Peri and Yasenov (2019); Clemens and Hunt (2019) for a discussion on the effects of Mariel Boatlift (Card, 1990), Dustmann et al. (2016) for a discussion on methodologies and Jaeger et al. (2018) for a discussion of the "past settlement" instrument.

of up to 82% in southeastern provinces near the Syrian border. I leverage this massive and spatially concentrated labor supply shock by comparing manufacturing firms in non-host regions that are differentially exposed to immigrants through their trading partners. I merge three administrative datasets—value-added-tax (VAT) records capturing the near-universe of firm-to-firm transactions, matched employer-employee records, and firm balance sheets—to calculate model-defined trade exposures for all formal firms in Turkey. To address endogeneity in refugees' location choices, I construct a shift-share instrument where the shift captures aggregate Syrian refugee inflows and the share reflects each region's travel distance from the Syrian border. This regional instrument translates to firm-level variation through baseline input-output linkages. I apply the Synthetic IV method (Gulek and Vives-i Bastida, 2024) to relax the standard share-exogeneity assumption (Goldsmith-Pinkham et al., 2020), allowing me to compare firms on similar trajectories while exploiting variation induced by the instrument for identification. I focus on firms with at least 50 employees to minimize the bias from informal labor and sales.[2]

The empirical results align with my theoretical mechanisms. First, firms that directly or indirectly buy from host regions increase their labor demand: they hire more workers and increase both payroll and the labor share in production costs. This implies that labor and intermediate goods are gross complements. Second, buyer firms maintain stable spending patterns across their suppliers on average, implying that intermediate good production approximates Cobb-Douglas. Third, firms that sell to host region firms increase their sales, which implies that intermediate goods are more substitutable with each other than with labor, a finding that reinforces the first two results. My regression specifications have direct structural interpretation, allowing me to recover the model's elasticities from the estimated coefficients. I use GMM to efficiently estimate the two elasticities, and find an elasticity of substitution between labor and intermediates of 0.79 and an elasticity across intermediates of 1.06. These estimates remain similar across robustness checks, and over-identification tests fail to reject the model.

Having estimated the key elasticities empirically, I turn to counterfactual analyses to understand for which types of host regions the spillover effects of immigration become economically significant. I simulate a 1% increase in labor supply for each of Turkey's 81 provinces separately and calculate the resulting changes in real wages across all regions. For 76 provinces, spillovers are negligible: a 1% increase in local labor supply reduces real wages by approximately 1% in the host region while increasing wages by less than 0.02% in non-host regions. However, immigration to central regions generates substantial spillovers. For instance, a 1% increase in Istanbul's labor supply reduces local real wages by only 0.56% while increasing real wages in the average non-host region by 0.38%, a spillover effect nearly two-thirds the magnitude of the direct effect. While both population size and economic development correlate with spillover magnitude, a region's centrality in the production network is the strongest predictor. Greater centrality flattens the labor demand curve in the host region and shifts it rightward in non-host regions, resulting in smaller wage decreases for natives in host regions and larger wage increases in non-host regions.

[2] Large firms being more formal than small firms is a common attribute of developing economies (Ulyssea, 2020).

I conduct a second counterfactual analysis that holds the absolute number of immigrants fixed across simulations, rather than fixing the immigrant-to-native ratio as in the first exercise. This alternative approach directly addresses a crucial policy question facing governments during refugee crises: how does the spatial allocation of immigrants affect aggregate welfare? I find that directing immigrants to economically central regions generates welfare gains that are an order of magnitude larger than placement in non-central regions. When immigrants settle in well-connected regions, their reduction of local production costs cascades throughout the economy through trade linkages, creating positive spillovers that can outweigh the negative wage effects on host region natives. The importance of network position extends to skill composition: high-skill immigration generates larger spillovers and, therefore, larger welfare gains than low-skill immigration because industries that employ high-skill labor intensively tend to have stronger inter-regional trade connections.

In the final analysis, I quantify the aggregate impact of Syrian immigration to Turkey by simulating a low-skill immigration shock that matches the observed spatial distribution of refugees. Because Syrians predominantly settled in non-central southeastern regions of Turkey, aggregate spillovers are negligible. The variation in regional wage effects is almost entirely explained by local immigrant-to-native ratios. While individual firms connected to host regions experience spillovers—as shown in my empirical results—these firm-level effects aggregate to economy-wide impacts only when immigrants settle in central nodes.

This paper contributes to the extensive empirical literature studying the economic effects of immigration.[3] A key methodological debate centers on whether immigration effects spillover beyond host regions. While Card (1990, 2001) assume effects are local and compare host to non-host regions, Borjas (2003) argues that spatial integration causes spillovers through native migration responses. However, empirical evidence shows that immigrants have minimal impact on native migration patterns (Card, 2009; Dustmann et al., 2017; Edo and Özgüzel, 2023), including in the present setting (Gulek, 2024).[4] My main contribution is demonstrating that immigrants impact non-host regions through input-output networks. I formalize the main economic forces by extending the production network framework (Acemoglu et al., 2012; Baqaee and Farhi, 2019), test them empirically using novel methods (Gulek and Vives-i Bastida, 2024), and quantify when general equilibrium effects become economically relevant.

My quantitative findings help resolve existing puzzles in the immigration literature. Despite 30 years of research following Card (1990), wage and employment effects of immigrants remain debated. My results suggest this reflects differences in the network centrality of host regions across studies. Consider Dustmann et al. (2017) and Beerli et al. (2021), two papers using identical designs but finding opposite results. Dustmann et al. (2017) study immigration to German border regions containing "various small but no large cities" and find strong negative effects on native em-

[3]Seminal papers include Card (1990, 2001); Borjas (2003); Ottaviano and Peri (2012). See Hanson (2009); Lewis and Peri (2015); Dustmann et al. (2016) for reviews.

[4]Monras (2020) is a notable exception, which shows that native migration responses in the US are significant. In the Turkish setting, Gulek (2024) shows that changes in in- and out-migration in response to Syrian immigration have been minimal.

ployment. In contrast, Beerli et al. (2021) examine immigration to Swiss border regions, including Basel, a major hub for pharmaceuticals and chemicals, and find positive effects on natives. My framework predicts negligible spillovers from small German regions but substantial spillovers from network central Basel, violating SUTVA and complicating interpretation. The positive Swiss effects could reflect either: (i) negative spillovers to non-host regions that make host regions look relatively better, or (ii) positive spillovers between host regions that dominate local negative effects.[5] Without knowing the network structure and elasticities, one cannot distinguish these mechanisms.[6] More generally, my framework helps explain why studies of economic migration often find weaker displacement effects than refugee studies.[7] Economic migrants typically settle in large, central cities (Albert and Monras, 2022; Mayors of Europe, 2019) where spillovers to other regions are substantial and likely violate SUTVA. In contrast, refugees often settle near borders of entry (Baez, 2011; Del Carpio and Wagner, 2015), typically non-central regions where spillovers are minimal and, therefore, SUTVA holds, allowing for clean identification of displacement effects.

A related literature examines the interaction between immigration effects and output tradability (Dustmann and Glitz, 2015) and international trade (Caliendo et al., 2021; Brinatti, 2024). Most notably, Burstein et al. (2020) formalize how industry tradability shapes local labor market responses to immigration. I extend their framework by demonstrating that production networks play a crucial role in these adjustments. My analysis shows that beyond industry tradability, the input-output linkages between industries have first-order effects on local labor markets.

This work also contributes to the growing literature on shock propagation through production networks. Theoretical work by Acemoglu et al. (2012, 2016b, 2017) and Baqaee and Farhi (2019) explores how microeconomic shocks can spread through input-output networks to generate aggregate fluctuations.[8] Empirical studies have documented this propagation for various economic shocks, including trade disruptions (Acemoglu et al., 2016a) and natural disasters (Barrot and Sauvagnat, 2016; Boehm et al., 2019; Carvalho et al., 2021). In the context of immigration, Akgündüz et al. (2024) provide the closest empirical analysis to mine, documenting positive spillovers on firms' sales and employment through direct trade linkages to regions hosting Syrian refugees in Turkey. I extend their analysis in three ways. First, I formalize the economic mechanisms through which immigration propagates via production networks. Second, I estimate the key elasticities governing these mechanisms: labor and intermediates are gross complements, while intermediates are independent—implying firms with lower costs cannot gain market share through price competition. Third, I quantify the total effects of immigration on wages and welfare across all regions, and show when the general equilibrium effects of immigration become economically significant.

[5] For example, if four host regions each receive identical shocks that reduce local wages by 1% but increase wages in other host regions by 0.5%, the average treatment effect would be positive despite local displacement.

[6] This is not a critique of Beerli et al. (2021)'s analysis, which carefully documents the positive effects. Rather, my framework highlights how network spillovers can contribute to such positive findings when immigration occurs in economically central regions.

[7] For a review of papers on largely economic immigration, see Dustmann et al. (2016). For papers analyzing refugee crises of the last decade and finding large displacement of natives, see Gulek (2024); Bahar et al. (2024).

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

2 Theory

This section formalizes how a decrease in wages due to immigration in one region can spill over to other regions through the production network, and develops structural equations that guide the empirical analysis.

2.1 Setup

The economy consists of N firms indexed by i, R regions indexed by r, where each region is endowed with L_r labor.[9] 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 and intermediate goods $\{x_{i,j}\}_{j=1}^{n}$ to minimize costs subject to a constant returns nested-CES technology:

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

where A_i is a Hicks-neutral productivity shifter, y_i is total output, p_j is the price of good j, L_i 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. I assume common EoS across firms in both the upper and lower nests: σ_u denotes the EoS between labor and intermediates, and σ_l is the EoS between different intermediates.[10] Constant returns to technology requires $\sum_j \alpha_{i,j} = 1$. I assume that firms have constant and exogenous markups denoted by μ. Therefore, firm i sets price $p_i = \mu_i C_i$, where C_i denotes the unit cost.

Final Demand

All final goods consumption, as well as the ownership of firms, is local. I assume a representative consumer in each region r, who optimizes her Cobb-Douglas utility subject to a budget constraint

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

[10] The common EoS assumption across firms simplifies the exposition, but can be relaxed. The empirical analysis relaxes this assumption by estimating heterogeneity across industries and finds limited heterogeneity.

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 inelastically 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 labor supply.

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.

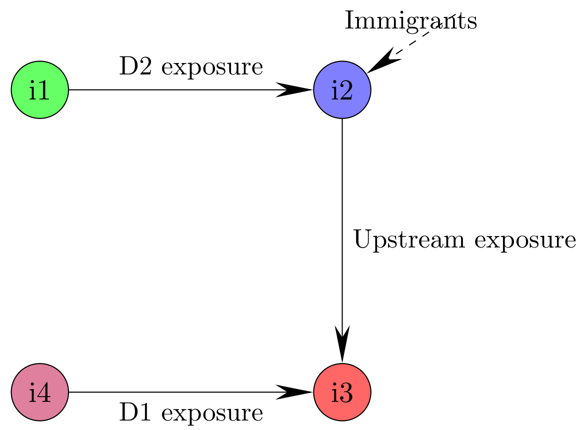
2.2 Three General Equilibrium Forces

The solution to this model is notation-heavy and, therefore, hard to follow. To facilitate exposition, I describe the three relevant economic forces here. Figure 1 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, a decrease in production costs decreases prices. This creates a chain reaction along the supply chain that propagates both forward and backward.

First, firm i_3 benefits from immigration as the price of the input from firm i_2 decreases. As i_3 faces lower input prices, it can increase or decrease its labor demand depending on the substitutability between intermediates and labor. If labor and intermediates are gross complements, then the reduction in input prices would cause firm i_3 to increase its labor demand. I name this the "upstream exposure effect" of immigration: upstream because the shock comes from upstream from the recipient i_3's perspective.

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 gross substitutes, then as i_2's prices decrease 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 occur: i_3 would increase its demand for i_4's goods, which would increase i_4's demand for labor. I name this the first downstream exposure effect, which I denote shortly as $D1$ for the rest of the paper.

Figure 1: 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.

Notice that the effects on both i_3 and i_4 are parts 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. First, the price of labor decreases compared to its input from i_1. The more substitutable the two inputs are, the less i_2 demands i_1's goods. Second, i_1 incurs a demand shock based on i_3's choice among goods from i_2 and i_4. More substitutable intermediates are, more i_3 demands from i_2, which results in i_2 demanding more from i_1. These two forces oppose each other. As I prove later, the net effect on i_1's sales depends on the relative magnitudes of the 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. I name this the second downstream exposure effect, which I denote shortly as $D2$ for the rest of the paper. This captures the backward propagation of the immigration shock.

Figure 1 only depicts the first-degree trade exposures: that is, firms being impacted by their immediate customers and suppliers. However, these forces expand beyond the first-degree linkages. Firms that indirectly buy from immigrant-intensive firms are also upstream exposed. The same applies to downstream exposures. 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, I use the model.

2.3 Input-Output definitions

To derive the impact of regional labor supply shocks on labor demand across all regions, I establish input-output notation following Baqaee and Farhi (2019).[11] My results are comparative statics describing how the labor payments in any host and non-host region change when a host region

[11] I maintain their notation except where my model's regional labor markets necessitate modifications.

receives immigrants. I 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 I present both. All these objects are defined at the initial equilibrium. Without loss of generality, I normalize the nominal GDP to 1. Finally, in my analytical results and counterfactuals, I assume constant markups and technology.[12]

2.3.1 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.

2.3.2 Input-Output Matrices

To streamline the exposition, I treat labor as a special endowment producer that does not use any input to produce. I 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, I 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 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

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

[12] 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).

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 are 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. I define the share of labor and intermediate good expenditures of firm i as:

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

2.3.3 Leontief Inverse Matrices

I 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.

2.3.4 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 the results, I 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, one 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.

2.4 Effects of a Labor Supply Shock on Labor Income

To build intuition as to how an immigration shock in a host region can impact the labor payments in any region, I 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, I 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:*

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

All proofs are in the Appendix.

Equation 1 captures the forward propagation of cost shocks, which is the upstream exposure effect introduced in Figure 1. Firms' labor share is determined by the trade-off firms face between using labor or intermediate goods in production. Suppose the local wages decrease 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.

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 gross substitutes, $\sigma_l > 1$, 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 who observe larger decreases in costs among the suppliers of firm j are also more dependent on firm i for production. Summing across all firm js 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, intermediate goods are more substitutable with each other than with 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 extent 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.[13]

Given the intuition developed in Propositions 1 and 2, I 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 is 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 firms gives 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 multiplied by the quantity of labor in that region: $\lambda_r = L_r w_r$. Combining this with Propositions 1, 2, 3, I can fully characterize the impact of immigration on this economy.

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

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

tities 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 I have described in one system of linear equations. Notice that I observe all the parameters in this equation in the pre-shock data except for the elasticity parameters σ_u and σ_l. 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 in this model.

2.5 Extensions and Limitations

The model formalizes how immigration shocks propagate through production networks to affect labor demand throughout the economy, but several limitations and extensions warrant discussion.

Labor mobility: The model assumes labor immobility across regions to isolate trade spillovers from migration responses. While native migration can equilibrate regional labor markets (Monras, 2020), this assumption holds empirically in the Turkish context: Syrian immigration induced minimal changes in native migration patterns (Gulek, 2024), which is also documented in Appendix Figure C.5.

Sufficient statistics: Theorem 1 does not yield simple sufficient statistics for predicting spillover magnitudes, limiting intuitive characterization of when general equilibrium effects deviate substantially from partial equilibrium predictions. I address this numerically through counterfactual analyses in Section 5.

Network formation: The model abstracts from firms' capacity to form new trading relationships. Immigration-induced wage reductions could prompt firms to establish new buyer-supplier relationships with host regions. While analyzing endogenous network formation exceeds this paper's scope, I address this empirically by exploiting the empirical fact that new trade formation predominantly occurs within regions or with firms maintaining existing linkages to host regions. Consequently, I construct firm-level exposure measures based on connections to region-industry cells rather than individual firms, as detailed in Appendix Section B.1. Counterfactual analyses assume representative firms at the region-industry level for theoretical clarity and computational tractability.

Skill heterogeneity: The model includes one skill type and assumes perfect substitutability between natives and immigrants. Both assumptions can be relaxed easily by specifying labor L as a CES aggregate of different worker types. Appendix Section A.3 presents this extension. Intuitively, when labor types are not perfect substitutes, the arrival of one type of labor creates an additional productivity shock from labor complementarity, but the main predictions of the model remain similar.

Labor supply elasticity: The model assumes perfectly inelastic labor supply in each region, abstracting from evidence that immigrants can displace natives. With elastic labor supply $L^S(w) = w^{\eta}$, $\eta \neq 0$, wage changes in host regions would be attenuated, limiting spillover magnitudes in non-host regions. Under a perfectly elastic labor supply ($\eta \to \infty$), spillovers would vanish completely since wages remain fixed, eliminating the production cost reductions that drive network propagation.

2.6 Structural Equations

Having solved the model, I now formalize the economic forces from Figure 1 and derive the structural equations that guide the empirical analysis. I introduce time subscript t and approximate wage changes in region r as $d\ln w_{rt} \approx \frac{\delta_{rt}}{\epsilon^D}$, where δ_{rt} denotes the immigration shock and ϵ^D denotes the elasticity of labor demand. Substituting this approximation into equations 1 and 2 yields:

$$
\begin{aligned}
d\ln\tilde{\Omega}_{i,L} &= -\frac{(1-\sigma_u)}{\epsilon^D} \underbrace{\sum_{r=1}^{R} \tilde{\Psi}_{i,r}\delta_{rt}}_{\text{Upstream Exposure (U)}} + \epsilon_{1,it} \\
d\ln\lambda_i &= \frac{(1-\sigma_l)}{\epsilon^D} \underbrace{\sum_{j=1}^{n} \frac{\lambda_j}{\lambda_i \mu_j} Cov_{\tilde{\Omega}(j)} \left(\sum_{r=1}^{R} \tilde{\Psi}_{(r)}\delta_{rt}, \Psi_{(i)} \right)}_{\text{Downstream-1 Exposure (D1)}} \\
&\quad + \frac{(\sigma_u-\sigma_l)}{\epsilon^D} \underbrace{\sum_{j=1}^{n} \frac{\lambda_j}{\lambda_i} \tilde{\Omega}_{j,l} \left(\delta_{r_j,t} - \sum_{k=1}^{n} \frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}} d\ln p_k \right) (\Psi_{ji} - I_{ji})}_{\text{Downstream-2 Exposure (D2)}} + \epsilon_{2,it},
\end{aligned}
\tag{5}
$$

where the error terms ϵ_1 and ϵ_2 capture the direct effects of immigration on host-region wages, spillovers of immigrant demand, and potential model misspecification arising from, for example, the constant markup and CES assumptions.

These equations reveal two key insights. First, the trade-exposure measures [U, D1, D2] depend only on the pre-shock trade network structure and the spatial distribution of immigration, making them directly computable from the data. Second, the coefficients on these exposure measures identify the structural elasticities: the upstream exposure coefficient recovers σ_u, while the downstream exposure coefficients jointly identify σ_l and confirm the ordering of elasticities. Therefore, the empirical analysis focuses on estimating these reduced-form relationships to recover the key EoS parameters that govern how immigration shocks propagate through the production network.

3 Background and Data

3.1 Syrian Refugee Crisis in Turkey

The Syrian Civil War, beginning in March 2011, displaced 6 million Syrians by 2017. Turkey absorbed the largest share—3.6 million registered refugees—dwarfing the numbers in neighboring Lebanon, Jordan, and Iraq combined. Figure 2a documents the rapid escalation: from 170,000 refugees in 2012 to 3.6 million by 2019, with the steepest increases occurring between 2013 and 2015.

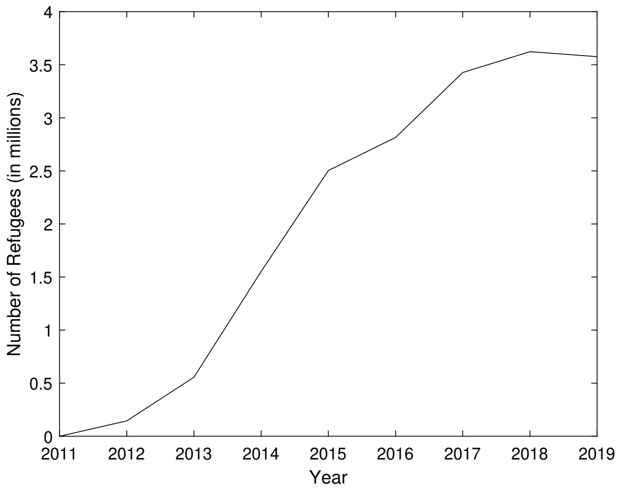
The Turkish government initially tried to host the Syrians in refugee camps in the southeastern part of the country across the Turkish-Syrian border. However, the camps quickly exceeded capacity as the number of arriving refugees increased. The refugees thus dispersed across Turkey.[14] Figure 2c shows the distribution of the number of Syrian refugees per 100 natives in Turkey at the province level. Refugees are more densely located in regions closer to the border. Distance to the populous governorates in Syria strongly predicts the number of refugees per native in a given region, which constitutes the backbone of the identification strategy.

Syrian refugees are less educated than Turkish natives. Figure 2b shows that 21% of Syrian refugees lack primary education compared to 12% of natives, and 83% do not have a high school diploma compared to 61% of natives. These educational gaps likely understate the effective skill differential. Educational downgrading is common among immigrants (Dustmann et al., 2013), and most Syrians possess only basic Turkish proficiency (Turkish Red Crescent and WFP, 2019), further limiting their access to skilled positions. The Syrian influx therefore constitutes a low-skill labor supply shock to the Turkish labor markets.

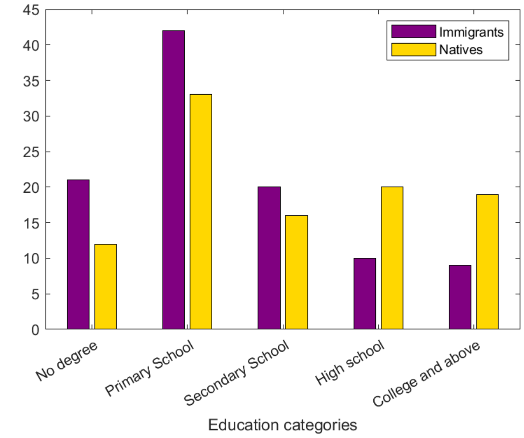
Most Syrians in Turkey do not have formal labor market access. As of March 2019, only 31,000 refugees—1.5% of the working-age population—held work permits, constraining the vast majority to informal sector. However, this does not undermine the generalizability of my findings. Gulek (2024) demonstrates that informal and formal labor are gross substitutes in production, with an elasticity of substitution of around 10. This high substitutability ensures that wage pressures from informal immigration transmit to the formal sector, generating an aggregate wage-elasticity of labor demand of approximately -1.27. This estimate allows me to quantify the reduction in labor costs

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

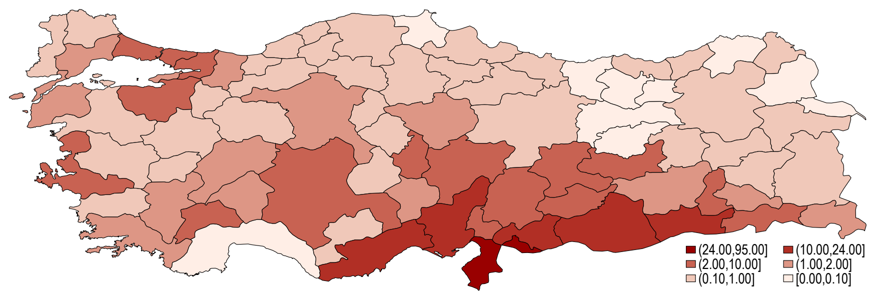
Figure 2: Statistics on the Syrian Refugees in Turkey



(a) Time series 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 comes from the Directorate General 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.

across host regions and trace how these cost shocks propagate through the production network to non-host regions.

3.2 Data

Studying the network spillovers of immigration shocks requires a comprehensive dataset covering who firms trade with, how much they spend on labor and intermediates, and how much they sell. To achieve this, I integrate five datasets covering all formal firms in Turkey between 2006–2019. The Ministry of Industry and Technology maintains these datasets with a unique and homogeneous firm identifier, which enables me to merge them.

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, I use the yearly gross sales of each firm. Third, from the firm registry, I 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, I collect firms' annual exports and imports. Fifth, from the employer-employee data, I collect the average number of workers, total labor costs, and average wages per worker for each year.

The network data are supplemented with labor force surveys conducted by the Turkish statistical institute. Unlike the census data, these surveys collect information on workers' education, which allows me to determine the skill intensity of industries and regions.

Data on the number of refugees in Turkey across years and provinces are acquired from the Directorate General of Migration Management of Turkey (DGMM). DGMM does not share the education and age breakdown of refugees at the province level, which prevents the empirical investigation from exploiting that variation.

Lastly, computing the trade exposure variables introduced in equation 5 requires substantial computational resources due to the need to invert large matrices. The baseline sample includes approximately 230,000 firms, generating trade matrices with 53 billion elements. While the trade matrices $\tilde{\Omega}$ and Ω remain sparse and computationally manageable, their Leontief inverses $\tilde{\Psi}$ and Ψ are dense and memory-intensive. To overcome this computational constraint, I provided a 512 GB RAM workstation to Turkey's Ministry of Industry and Technology. Appendix Section B.1 provides detailed documentation of the matrix construction and treatment variable calculations.

4 Empirical Analysis

This section presents evidence on how immigration-induced trade spillovers affect manufacturing firms in non-host regions. I describe the identification strategy, estimate the causal effect of trade exposures on firms' labor demand and sales, estimate the key EoS parameters, and use these parameters to quantify the total equilibrium effects of immigration across all regions.

4.1 Identification Strategy

There are three threats to identification. First, trade exposures depend on regional immigration intensities (δ_{rt}), which may be endogenous if immigrants select regions with positive labor demand shocks. Second, trade exposures depend on input-output matrices (Ω and $\tilde{\Omega}$), creating potential bias if firms with different trade exposures follow different trajectories. Third, Turkey has a large informal sector: 40% of employment and an unknown share of sales remain unregistered and therefore absent from the census data. This mismeasurement affects both outcome and treatment variables—the latter non-linearly—potentially biasing estimates in either direction.

I address these three challenges through a corresponding three-step approach. I use shift-share instrumental variables (SSIV) to tackle the endogeneity concern, combine this with Synthetic Controls (SC) to address the differential trajectories concern, and focus on large firms that are substantially more formal to mitigate the informality concern. I first introduce the instruments, then explain how I utilize the Synthetic Instrumental Variables (SIV) approach of Gulek and Vives-i Bastida (2024) to combine instrumental variables for immigration patterns with synthetic controls for firm trajectories, before finally explaining why focusing on large firms shields the analysis from bias from informality.

4.1.1 Instruments

My primary shift-share instrument for immigrant location choices combines inverse travel distances between Turkish regions and Syrian governorates (share) with the total Syrian refugee population in Turkey (shift):

$$Z_{r,t} = \underbrace{\sum_{s=1}^{13} \lambda_s \frac{1}{d_{r,s}}}_{\text{Share}} \times \underbrace{\text{Number of Syrians in Turkey in year } t}_{\text{Shift}}, \tag{6}$$

where $d_{r,s}$ measures travel distance between region r and governorate s, and λ_s weights each governorate.[15] Following Aksu et al. (2022), I weight governorates by their population and proximity to Turkey relative to other neighboring countries,

$$\lambda_s = \underbrace{\frac{\frac{1}{d_{s,T}}}{\frac{1}{d_{s,T}} + \frac{1}{d_{s,L}} + \frac{1}{d_{s,J}} + \frac{1}{d_{s,I}}}}_{\text{Relative distance to Turkey}} \times \underbrace{\quad \pi_s \quad}_{\text{Pop. share}}, \tag{7}$$

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

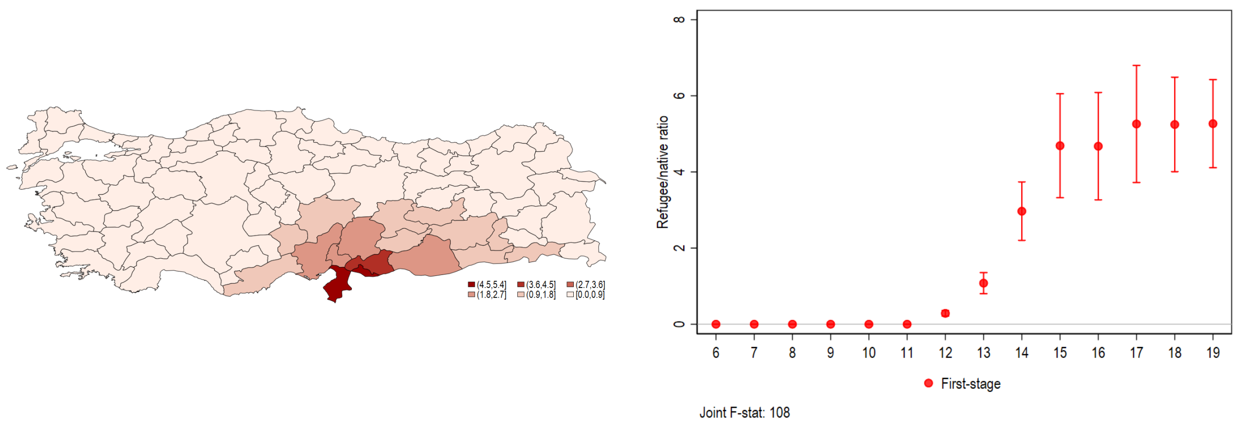
Figure 3a shows the cross-sectional distribution of the distance share component of the instrument. It puts higher weights in southeastern Turkey near northwestern Syria, reflecting the higher Syrian population density around Aleppo (northwest of Syria) compared to Al-Hasakah (northeast of Syria) along the Turkish border. Figure 3b shows the first-stage estimates from a nonparametric event-study design where I regress the immigration treatment δ_{rt} on the distance-share Z_r interacted with year indicators. Estimates between 2006–2011 are zero as there are practically no Syrian immigrants in Turkey during those years.[16] After 2011, distance strongly predicts immigrant location in all years. The joint F-statistic for the post-period coefficients is 108, which implies a strong instrument.

I validate the main instrument with two alternative shares. The first one uses the share of Arabic speakers from the 1965 census. Unlike Card (2001)'s past-settlement instrument, the Arabic-speaking population reflects Ottoman Empire demographics rather than previous Syrian migration. The second alternative instrument uses a dummy indicator for the regions that receive large weights

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

[16] Although Syrians started arriving to Turkey in 2011, the initial numbers were minuscule.

Figure 3: Distance instrument distribution and first-stage



(a) Distance instrument 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 I weight each region by its population in 2011. Standard errors are clustered at the region level. 95% confidence intervals are plotted.

by either the distance or the language instruments.[17] This primarily serves as a sanity check as it is arguably easier to interpret reduced-form effects with a dummy indicator. The trade exposure instruments (U^z, $D1^z$, and $D2^z$) are constructed by replacing the regional immigration δ_{rt} with the regional instrument Z_{rt} in the respective exposure measures. The respective first-stage estimates of these alternative shares can be found in the Appendix Figure C.6.

4.1.2 Synthetic IV Design

The ideal comparison would contrast similar firms that would have followed parallel trajectories absent the immigration shock but happen to be differentially exposed to immigrants through their trading partners. However, as Appendix Section C demonstrates, firms with different trade exposures follow divergent pre-period trajectories even within narrow region-industry or region-industry-size cells. This divergence partly reflects stronger employment growth in southeastern Turkey during 2006–2011 (Gulek, 2024), which likely propagated through production networks to differentially affect firms in non-host regions. These pre-existing differences violate the share-exogeneity assumption underlying standard shift-share designs (Goldsmith-Pinkham et al., 2020).

To address this challenge, I implement the SIV estimator (Gulek and Vives-i Bastida, 2024), which proceeds in two steps. First, I construct synthetic controls for each firm using pre-period data and generate counterfactual values for outcomes, treatments, and instruments. Second, I apply standard IV estimation to debiased data, defined as the difference between observed and synthetic values.

[17] This coincides with the host regions shown in Figure B.1.

Formally, let ω_i denote the vector of synthetic control weights for firm i and X represent an $N \times K$ matrix containing outcomes, treatments, and instruments. Using the SC weights, I generate synthetic values $\hat{X}_k = \sum_{j \neq i} \omega_j X_{jk}$ for each variable k. I then subtract these synthetic values from the observed data to obtain *debiased* values: $\tilde{X} = X - \hat{X}$. This procedure effectively partials out unobserved confounders proxied during the matching step. Finally, I apply two-stage least squares to the debiased data. The resulting SIV estimator remains consistent even when standard IV fails (Gulek and Vives-i Bastida, 2024).

To construct synthetic control weights, I match on pre-2011 demeaned values of log labor share and log sales. Following Sun et al. (2025), I estimate a single set of weights for both outcomes, which improves signal-to-noise ratios. Appendix Section D confirms that joint estimation outperforms separate weights when predicting unmatched outcomes like payroll and firm size. I restrict donor pools to firms within the same region and two-digit industry, effectively controlling for region-by-industry-by-time fixed effects. Following Abadie and L'hour (2021), I include a penalty term to mitigate overfitting in this disaggregated setting.

This approach compares firms within region-industry cells that *followed similar economic trajectories before the immigration shock* but experienced different exposure to immigrants through their trading networks.

4.1.3 Sample Selection

To address bias from informality, I exploit the well-documented inverse relationship between firm size and informality rates (Ulyssea, 2018, 2020). Appendix Figure B.3 confirms this pattern in Turkey: informality rates decline sharply with firm size, from 60% in firms with fewer than 10 employees to below 5% in firms exceeding 50 employees. The 50-employee threshold marks a particularly sharp discontinuity, as it triggers heightened regulatory scrutiny that plausibly reduces informality in both employment and sales reporting. I therefore restrict the sample to firms with at least 50 formal employees in 2010. Appendix Section D.5 demonstrates robustness to alternative size thresholds.

Two additional restrictions sharpen the identification. First, I exclude firms in regions where the immigrant share exceeds 4% of the native population or where the instrument assigns large weights, ensuring that estimated effects operate through trade linkages rather than the direct impacts of immigration. Appendix Figure B.1 illustrates the excluded regions. Second, since SIV estimator requires a balanced panel, I restrict the sample to firms with non-missing observations in employment, wage bill, and sales between 2006–2019. I also focus exclusively on manufacturing firms in the empirical analysis. There are 19,155 such firms, 1,224 of which had 50+ employees in 2010. Summary statistics are reported in Appendix Table B.1.

4.2 Estimating Equations

4.2.1 Synthetic IV Design

Given trade exposure treatments U, $D1$, $D2$, and their respective instruments U^z, $D1^z$, and $D2^z$, I define the estimating equations of the structural equation 5 as follows:

$$\begin{aligned} \widetilde{log(LaborShare_{it})} &= \beta_1 \widetilde{U_{it}} + f_i^L + f_t^L + \nu_{it}^L \\ \widetilde{U_{it}} &= \gamma_1 \widetilde{Z_{it}^U} + g_i^L + g_t^L \omega_{it}^L \end{aligned} \tag{8}$$

$$\begin{aligned} \widetilde{log(Sales_{it})} &= \beta_2 \widetilde{D1_{it}} + \beta_3 \widetilde{D2_{it}} + f_i^S + f_t^S + \nu_{it}^S \\ \widetilde{D1_{it}} &= \gamma_2 \widetilde{Z1_{it}^D} + \gamma_3 \widetilde{Z2_{it}^D} + g_i^S + g_t^S + \omega_{1,it}^S \\ \widetilde{D2_{it}} &= \gamma_4 \widetilde{Z1_{it}^D} + \gamma_5 \widetilde{Z2_{it}^D} + h_i^S + h_t^S + \omega_{2,it}^S \end{aligned} \tag{9}$$

where $\tilde{X}$ refers to the debiased version of the variable X following the SIV algorithm.

Two important considerations guide my specification choices. First, equations 8 and 9 reflect the correct structural relationships for identifying elasticity parameters, which explains the separate treatment of upstream and downstream exposures. Second, the upstream exposure measure U is measured with greater precision than the downstream exposures $D1$ and $D2$ because sales are noisier than costs. Including U in equation 9, while theoretically unnecessary, could capture the causal effects of the noisier downstream measures in a joint estimation. Nevertheless, Appendix Section C demonstrates that the main findings remain robust when estimating upstream and downstream effects simultaneously.

4.2.2 GMM Design for Elasticity Estimation

The reduced-form coefficients from equations 8 and 9 map to the structural elasticities through the following relationships:

$$\beta_1 = -\frac{(1-\sigma_u)}{\epsilon^D} \quad ; \quad \beta_2 = \frac{(1-\sigma_l)}{\epsilon^D} \quad ; \quad \beta_3 = -\frac{(\sigma_l - \sigma_u)}{\epsilon^D}, \tag{10}$$

where $\epsilon^D = -1.27$ is the wage elasticity of labor demand from Gulek (2024). With three reduced-form coefficients identifying two structural parameters, the system is overidentified. This overidentification provides a specification test: if the model is correctly specified, all three coefficients should yield consistent elasticity estimates.

I estimate σ_u and σ_l via GMM. Define $\hat{\beta} \equiv [\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3]'$ as the vector of reduced-form estimates with covariance matrix $\hat{\Sigma}$, and let $\beta(\sigma)$ denote the theoretical mapping in equation 10. The GMM estimator solves:

$$\min_{\sigma_u, \sigma_l} J \equiv (\hat{\beta} - \beta(\sigma))' \hat{\Sigma}^{-1} (\hat{\beta} - \beta(\sigma)). \tag{11}$$

The minimized objective function provides Hansen's J-statistic for testing overidentifying restrictions. Under the null hypothesis that the model is correctly specified, $J \sim \chi^2(1)$ with one degree of freedom. I reject the model specification if J exceeds 3.84, the 5% critical value.

4.2.3 Event-study Design

The primary advantage of the event-study design is that it allows me to visually and flexibly assess the pattern of outcomes the *debiased* share component of the shift-share instrument captures relative to the beginning of the refugee crisis. The event-study equations of the SIV estimator for labor share are defined 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.

4.3 Empirical Estimates

4.3.1 Effects of Upstream and Downstream Exposures

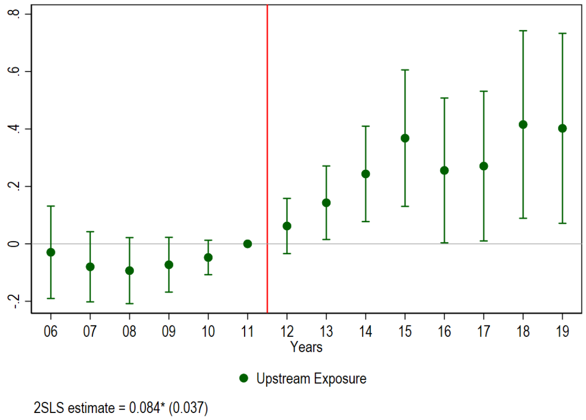
I begin by estimating the reduced-form effects of upstream exposure on firms' labor demand. Figure 4 plots the results for three outcomes: number of employees (Panel A), total payroll (Panel B), and labor share (Panel C). Panel D shows the effects of downstream exposures on sales. For inference, I cluster standard errors at the firm level.

Figure 4a reveals two main patterns. First, there are no statistically or economically significant pre-trends. This is not mechanical since SIV weights are generated to match trends in labor share and sales, not payroll or firm size. The absence of pre-trends in Figure 4a, despite the raw IV design showing pre-trends (as demonstrated in Online Appendix Section D.1), provides evidence in favor of the identification strategy. It shows that an unobserved common factor generates differential trends between more- and less-exposed firms, and that SIV successfully partials out this confounder.

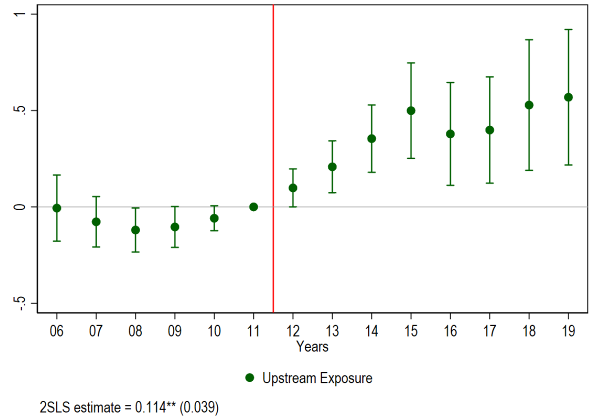
Second, upstream exposure causes firms to expand employment. Firms in non-host regions that directly or indirectly buy from immigrant-intensive firms in host regions hire more workers. The estimated effects grow over time, paralleling the pattern in the first-stage results, a similarity that strengthens the causal interpretation of these findings.

Interpreting coefficients from this reduced-form design requires careful consideration of how general equilibrium exposures propagate. Consider a simple example with two firms, i_1 and i_2, where each spends half its costs on labor and half on one intermediate good, buys from different suppliers (j_1 and j_2 respectively), and has suppliers that also allocate half their costs to labor. If firm j_1 has two standard deviations higher immigrant exposure through distance than firm j_2, this creates a 0.5 unit difference in upstream exposure between their customers i_1 and i_2. The 0.22

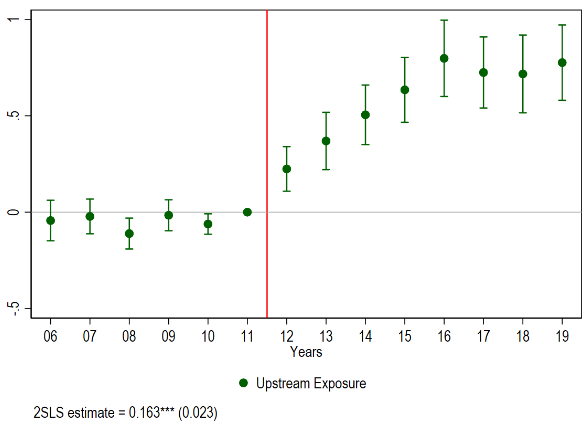
Figure 4: Effects of trade exposures on firms' labor demand and sales



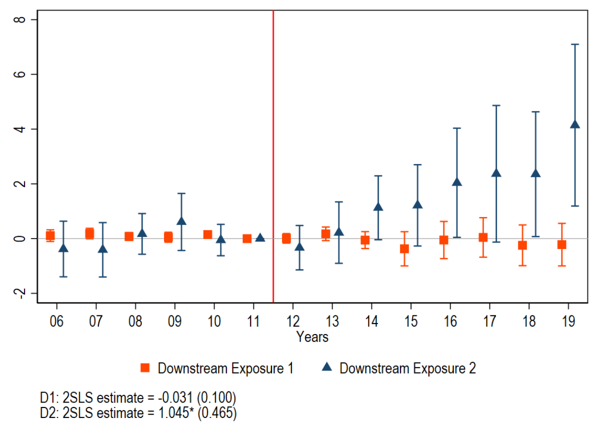
(a) Number of employees



(b) Payroll



(c) Labor share



(d) Sales

Notes: In Panels A, B, and C, 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. In Panel D, the estimates come from the regression equation: $\widetilde{y_{it}} = \sum_{t' \neq 2011} \left(\gamma_{2,t'} \widetilde{D1_i^Z} + \gamma_{3,t'} \widetilde{D2_i^Z} \right) \mathbb{1}\{t = t'\} + f_i + f_t + \nu_{it}$. Both the outcome and the treatment are their debiased versions following the SIV algorithm. The sample is restricted to manufacturing firms with at least 50 employees in 2010. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

coefficient estimated for 2019 in Panel A thus implies that firm i_1 increases its size by 11% relative to firm i_2.

Figure 4b presents the effects of upstream exposure on payroll. The effects parallel those for employment: no significant pre-trends and a positive, increasing difference between more- and less-exposed firms. The estimated effects on payroll modestly exceed those on employment, indicating that upstream exposure weakly increases wages, although this wage effect is not statistically significant.

Figure 4c shows the effects of upstream exposure on firms' labor share. The absence of pre-trends during 2006–2011 demonstrates good pre-treatment fit in the training period, a crucial condition for SIV validity since labor share is included in the matching step. Starting in 2012, upstream-exposed firms show significant increases in labor share: firms in non-host regions that directly or indirectly buy from host regions increase their labor share relative to similar firms in their region-industry cells. The 2SLS estimate, reported in the figure, is statistically significant (p-value <0.001). This result implies that labor and intermediate goods are gross complements.

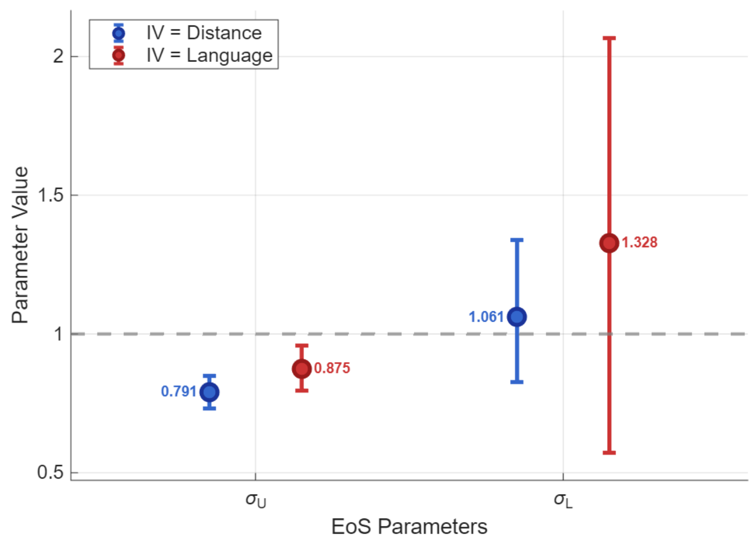
Figure 4d shows the effects of downstream exposure on firms' sales. I observe no significant pre-trend between 2006–2011. In the post-period, while I document precise null effects of $D1$ exposure on sales (p-value: 0.76), $D2$ exposure significantly increases firms' sales (p-value: 0.025). This pattern yields two key insights. First, consider a firm with suppliers from both host and non-host regions. The null $D1$ effect implies that this firm maintains fixed spending across suppliers and, therefore, does not substitute across them. This pattern is consistent with intermediate good production being Cobb-Douglas. Second, the positive $D2$ effect implies that firms selling directly or indirectly to host regions increase their sales, which occurs only if intermediate goods are more substitutable with each other than with labor. This result is consistent with the positive upstream exposure effect on labor share (implying labor and intermediate goods are gross complements) and the null $D1$ exposure effect on sales (implying intermediate goods are independent). These mutually consistent estimates provide an intuitive, albeit informal, test of the model.

To estimate the structural elasticity parameters efficiently and provide a formal test of the model, I bootstrap equations 8 and 9 using 1,000 draws with replacement to obtain the covariance matrix of regression parameters $[\beta_1, \beta_2, \beta_3]$. Then, I solve equation 11 to obtain GMM estimates of structural elasticity parameters. Figure 5 shows the results. Using the distance instrument, I estimate that labor and intermediates are gross complements, with an EoS of around 0.79, and intermediate goods are independent, with an estimated EoS of around 1.06. I can statistically conclude that labor and intermediates are gross complements since the estimated elasticity is statistically less than 1, and I cannot statistically reject the hypothesis that intermediate goods are independent.

These conclusions remain robust when using the language-based instrument. The language instrument predicts slightly larger substitutability for both technologies, but with lower precision because of weaker first-stage results. I cannot statistically distinguish estimates from distance and language instruments. Furthermore, overidentification tests do not reject the model using either instrument, increasing confidence in the validity of the results.

To summarize, Figure 4 shows that upstream exposure increases firms' labor demand, $D1$ exposure has a null effect on sales, and $D2$ exposure increases sales. This means that firms buying from host regions increase their labor demand, firms selling to host regions increase their sales, and firms do not substitute across suppliers on average. Using the preferred distance-based instrument, I estimate the elasticity parameters as: $[\hat{\sigma}_u = 0.79, \hat{\sigma}_l = 1.06]$.

Figure 5: Elasticity of substitution estimates



This figure shows the 95% confidence intervals of the EoS estimates using two alternative instruments. I draw 1000 bootstrap samples with replacement from the sample of large manufacturing firms. I then estimate the effects of upstream exposure on labor share and downstream exposures on firms' sales. The joint covariance of the regression parameters from different regressions is calculated by using the bootstrapped distribution of the regression coefficients. 95% confidence intervals are displayed.

4.3.2 Robustness Checks

I perform extensive robustness checks to ensure that estimated effects represent causal impacts rather than differential trends. I report detailed results in the Appendix and summarize key findings here.

Alternative Instruments

Appendix Figure C.7 replicates the results for labor share and sales using travel distance, language share, a dummy indicator for the host regions, and each share interacted with the skill-intensity of industries. The latter exploits the fact that Syrian immigrants are predominantly lower-skilled than Turkish workers and, therefore, overrepresented in low-skill-intensive industries. All six weighting schemes yield identical conclusions: upstream exposure significantly increases labor demand, $D1$ exposure has a null effect, and $D2$ exposure increases sales. The only caveat is that the language instrument is less precise, so language- or language-skill-based instruments do not yield statistically significant effects on sales. The results for labor share remain robust, since cost-based measures have much lower variance than sales-based measures.

Industry Heterogeneity

While the model allows firm-level heterogeneity, I assume constant elasticity of substitution parameters for computational reasons, yielding average estimates across industries. To relax this assumption, I estimate the elasticity parameters separately for each two-digit manufacturing industry, apply Empirical Bayes shrinkage to adjust for small sample bias, and plot results in Figure C.4. Overall, across different specifications, labor and intermediates are gross complements in most industries. Although intermediate goods are independent in most industries, they are gross substitutes in some (such as textiles and basic metals) and gross complements in others (such as chemicals). Details of industry-level estimation appear in the Appendix Section C.2.

SIV Validation

Since SIV is a novel estimator that has not yet been used widely, demonstrating its robustness is particularly important. SIV is a synthetic control-based estimator. As is common in SC estimation, two key concerns are under-fitting and over-fitting. Under-fitting occurs when no convex combination of donor units can match treated units, while over-fitting occurs when synthetic control weights match noise rather than signal. My estimator does not suffer from these concerns. The absence of pre-trends in targeted outcomes (labor share and sales) demonstrates that more exposed firms are not outliers: I successfully construct synthetic firms with similar trends. Furthermore, the lack of pre-trends in untargeted outcomes (firm size and payroll) provides evidence against over-fitting, since these variables were not used in calculating synthetic control weights.

In addition, I perform several checks to demonstrate the robustness to these concerns. First, I document why employing SC is necessary. Figure D.8 compares event-study estimates from IV and SIV designs. IV shows significant pre-trends in most variables. More upstream-exposed firms follow different economic trajectories in employment, labor costs, and labor share between 2006–2011 than less-exposed firms, even when comparing firms within the same region-industry with similar baseline sizes. This demonstrates the need for adjusting for pre-trends to obtain credible estimates.

I continue by showing the importance of matching on labor share and sales jointly rather than separately. The intuition is that with a limited training period, more signal can be obtained by matching on multiple key outcomes. Figure D.9 shows that pre-trends in untargeted outcomes (labor costs and employment) remain large and significant when matching separately on labor share and sales.

One advantage of matching separately on labor share and sales is that it improves the pre-treatment fit for each variable as an outcome. When I study the effects on labor share, matching on labor share only results in a better pre-treatment fit than matching on labor share and sales jointly. Figure D.10 shows that my main results remain robust to matching for each outcome separately. This robustness is expected since the main design shows no pre-trends, already providing evidence against under-fitting.

Results also remain robust to back-testing. Figure D.11 replicates the analysis when matching on trends during 2006–2009 and 2006–2010 instead of throughout 2006–2011. All results remain robust.

Firm Size Sensitivity

The remaining researcher choice not yet tested is the firm-size restriction. I focus on large firms to address potential concerns regarding informality in both labor costs and sales. The 50-employee threshold was chosen as a legal threshold above which companies face additional liabilities and reporting requirements, making their data more credible. Figure D.12 shows estimates of the effects of trade exposures on labor share and sales across firms of different sizes. While upstream exposure effects on labor share are similar across firm sizes, downstream exposure effects on sales (specifically, $D2$ exposure effects) are sensitive to the 50+ threshold, above which firms' reporting practices become significantly more trustworthy. Above the 50+ threshold, results are consistent across firms with 60+, 70+, 80+, 90+, or 100+ employees. Online Appendix Section D explains how small firm informality can bias $D2$ exposure effects.

4.3.3 Comprehensive Robustness Tables

Finally, I present extensive robustness checks where I (i) estimate upstream and downstream exposure effects on each outcome, (ii) estimate trade exposures separately and jointly, (iii) use all manufacturing firms and firms with 50+ employees, and (iv) use alternative instruments. Tables D.3 through D.8 show these comprehensive robustness checks, and the Online Appendix Section D.5 interprets the results in the specifications. Overall, my main results remain robust.

5 Counterfactuals

This section uses the model to quantify how immigration affects host and non-host regions through counterfactuals. I examine the economic significance of trade spillovers, their dependence on host region and immigrant characteristics, and their implications for both policymakers and researchers.

Theorem 1 characterizes immigration's general equilibrium effects on regional wages and firm prices as the solution of a system of linear equations that depend on the baseline production network and structural elasticity parameters. Having observed the network in the data and estimated the elasticities, I can solve this system to obtain the general equilibrium effects. For computational feasibility, I use representative firms at the region-industry level. I begin with homogeneous labor within regions as in Section 2, then introduce skill heterogeneity to analyze how the effects vary with immigrant skill levels.

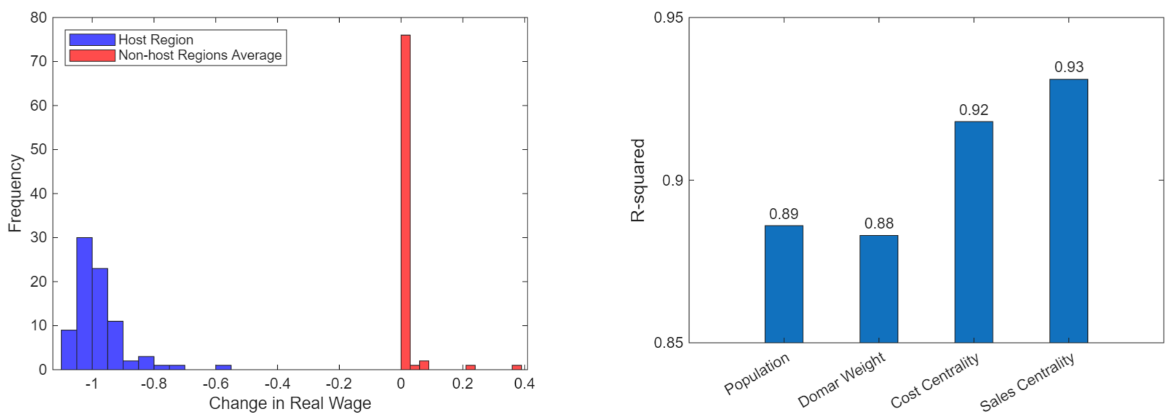
An important consideration is that the model expresses wages relative to nominal GDP, while real wages typically reference local prices. Therefore, I define real wage changes as $d\ln w_{real} = d\ln w - b * d\ln p$, where b denotes the $R \times N$ matrix of final expenditure shares and $d\ln p$ is the $N \times 1$ vector of price changes.

5.1 Spillover Effects of a 1% Labor Supply Shock

In the first counterfactual, I analyze how immigration spillovers vary across potential host regions. I simulate a 1% increase in labor supply separately for each of Turkey's 81 provinces and calculate two effects: the real wage change in the host province and the average real wage change in the other 80 provinces. This generates 81 pairs of estimates for host and non-host wage effects.

Figure 6a presents the distribution of wage effects, revealing two key patterns. First, a 1% increase in labor supply typically reduces the real wages of the host region by about 1% while leaving non-host regions largely unaffected. In 71 of 81 simulations, average non-host region real wages change by less than 0.01%, and in 76 cases by less than 0.02%. This pattern emerges because most firms predominantly trade within their own region, so host region price changes rarely generate economically meaningful spillovers to non-host regions.

Figure 6: Real wage changes in host and average non-host region



(a) Histogram of host and non-host region effects

(b) R-squared from bivariate regressions

(c) Heatmap of non-host region effects

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.

Second, 5 of 81 provinces generate economically meaningful spillovers (greater than 0.04% change in non-host real wages): Bursa, Kocaeli, Izmir, Ankara, and Istanbul. Istanbul and Ankara produce particularly large spillovers, up to two-thirds the magnitude of direct effects. A 1% labor supply increase in Istanbul reduces local real wages by 0.56% while raising the average non-host region's real wages by 0.38%. Similarly, in Ankara, a 1% shock decreases local real wages by 0.71% and increases average non-host wages by 0.22%. Figure 6c maps these spillover effects across regions. While the largest spillovers come from the most populated cities (Istanbul and Ankara), significant effects also emerge from major agricultural hubs (Manisa and Adana) and resource centers (Zonguldak with its coal deposits).

What explains this variation in regional spillovers? Population offers one explanation: a 1% labor force increase in Istanbul represents seven times the absolute immigration shock of a similar percentage increase in Gaziantep, a major host region. However, population alone cannot explain the pattern. Kocaeli, despite its smaller population than major host regions like Gaziantep, Sanliurfa, and Adana, generates larger spillovers than all three combined. Similarly, while Domar weights correlate with spillover magnitude, they don't tell the complete story. Consider Adana and Antalya: despite similar populations and Domar weights, Adana's spillovers are triple those of Antalya. This difference likely stems from their economic structures: Adana's role as an agricultural hub involves extensive inter-regional trade, while Antalya's tourism-focused economy generates mainly local transactions. This suggests that a region's position in the domestic trade network might better predict spillover effects.

I formally test this network position hypothesis using Bonacich centrality measures for both cost-based and sales-based trade matrices: $\widetilde{B} = \tilde{\Psi}'\mathbf{1}$ and $B = \Psi'\mathbf{1}$.[18] These measures capture how much other regions depend on region r through cost linkages ($\widetilde{B}_r$) and sales linkages (B_r). To assess which regional characteristics best predict spillover effects, I regress average non-host wage effects on the host region's population, Domar weight, and both centrality measures. Figure 6b presents the R-squared from these bivariate regressions. Sales-based centrality provides the strongest explanatory power (R-squared of 0.93 across 81 observations), confirming that a region's position in the production network determines spillover magnitude. Population shares, an easily observable statistic for regions worldwide, also predict spillovers well (R-squared of 0.89). This establishes a general empirical rule: immigration to large cities generates substantial spillovers to non-host regions.

Implications for Spatial Designs

Given that immigration to central cities generates sizable spillovers affecting labor demand in other regions, what implications does this have for spatial difference-in-differences designs commonly used in immigration research? To answer this question, I conduct one million simulations where I randomly assign 10 of 81 regions to receive a 1% immigration shock and calculate the average treatment effect on the treated (ATT) and control (ATC) regions. Under SUTVA, ATC would be zero, and therefore comparing the treated to control regions would provide a consistent estimate of ATT.

[18] For more on the Bonacich centrality measure, see Bonacich (1987) and Jackson (2008).

Appendix Figure E.14 presents the relationship between ATT and ATC under the estimated elasticity parameters ($\sigma_u = 0.79$, $\sigma_L = 1.06$). Three patterns emerge. First, the ATT on real wages varies substantially across simulations, ranging from -1.04% to -0.15%. This wide dispersion occurs because the composition of treated regions matters: when multiple central regions receive immigrants simultaneously, their positive spillovers on each other attenuate the direct negative effects. For example, a 1% immigration shock to Istanbul alone reduces local real wages by 0.56% while increasing Ankara's real wages by 0.35%. The same shock to Ankara reduces local real wages by 0.71% while raising Istanbul's real wages by 0.20%. When both receive immigrants simultaneously, their real wages decrease by only 0.36%—substantially less than the -0.64% average of their separate effects. Moreover, ATT can turn positive if accumulated spillovers are large enough: a 3% immigration shock to Istanbul and Ankara fully offsets the negative effects of a 1% shock to the four least central cities, yielding positive average real wage effects in host regions.

Second, ATC estimates are systematically positive, violating the stable unit treatment value assumption. When treated regions are predominantly central cities, control regions experience real wage increases up to 0.82%, making them inappropriate counterfactuals for estimating treatment effects.

Third, the relationship between ATT and ATC depends critically on production technology and the outcome variable. Figure E.14 Panels C and D show results for nominal wages—wages unadjusted for local price changes—a standard outcome in empirical studies. When labor and intermediates are gross substitutes ($\sigma_u = 10$, $\sigma_L = 1$), spillovers reverse direction: immigration to central cities reduces nominal wages in control regions through substitution toward intermediate inputs.

These simulations demonstrate that spatial difference-in-differences designs can yield misleading inferences when host regions are central nodes. A researcher finding positive employment effects or real wage effects might conclude that immigrants and natives are gross complements, when the result actually reflects positive spillovers from central to peripheral regions. Conversely, under different production technologies, the same design could both overstate or understate negative effects on nominal wages. The bias direction and magnitude depend on both network structure and technological parameters—neither of which is typically known to researchers ex ante.

These results shed new light on why immigration studies often reach conflicting conclusions, whether studying the same setting with different methods (Card, 1990; Borjas, 2017; Peri and Yasenov, 2019) or using identical designs across different settings (Dustmann et al., 2017; Beerli et al., 2021). Previous explanations include skill downgrading bias in skill-cell approaches (Dustmann et al., 2016), temporal confounding in past-settlement instruments (Jaeger et al., 2018), and sample composition effects (Clemens and Hunt, 2019). While these factors resolve part of the puzzle, they cannot fully explain persistent disagreements.

Consider two well-published papers using identical spatial difference-in-differences designs to exploit policy-induced immigration to border regions: Dustmann et al. (2017) study Czech immigration to Germany, while Beerli et al. (2021) examine EU immigration to Switzerland. Despite methodological similarity, Dustmann et al. (2017) find strong negative effects on native employment and wages, whereas Beerli et al. (2021) find positive effects. My results suggest an explanation beyond standard factors like labor market tightness or skill complementarity. The German host regions contain "various small but no large cities" (Dustmann et al., 2017, p.13), whereas the Swiss host regions include Basel, a major hub for pharmaceuticals, chemicals, and intermediate manufacturing. My framework predicts negligible spillovers in the former case but substantial positive spillovers in the latter, particularly across Switzerland's 2,000+ municipalities. Immigration-induced cost reductions in Basel likely increased labor demand in nearby host regions through domestic trade linkages, causing average host regions to experience more positive shifts than non-host regions. This mechanism can reconcile the conflicting findings across these otherwise comparable studies.

My numerical results also provide an explanation as to why different papers disagree on the effects of the Mariel Boatlift on Miami's labor market. In his seminal paper, Card (1990) examined the Mariel Boatlift's impact on Miami's labor markets by comparing Miami to Atlanta, Houston, Los Angeles, and Tampa and found null effects. These results are later challenged by Borjas and Monras (2017), who find large negative effects on low-skill natives' wages. This challenge itself is later challenged by Peri and Yasenov (2019), who use synthetic controls to address pre-trends and replicate Card's main findings. Clemens and Hunt (2019) show that contradictory findings on the effects of the Mariel Boatlift can be explained by a large difference in the pre- and post-Boatlift racial composition in certain very small subsamples of workers in the Current Population Survey.

My framework offers two complementary explanations for these patterns. First, if Miami maintains sufficient trade connections to the broader US economy, immigration-induced cost reductions would propagate through production networks, flattening Miami's labor demand curve. This mechanism, analogous to Istanbul's experience in my Turkish analysis, could generate genuinely small local wage effects despite substantial immigration inflows. The absence of detectable wage differences between Miami and comparison cities would then reflect economic reality rather than measurement error.

Second, even if wages in Miami decreased, network spillovers could simultaneously reduce labor demand in control regions, masking Miami's true treatment effects. This can happen if firms in the control regions directly or indirectly buy from firms in Miami, and labor and intermediates are gross substitutes for the main industries in these regions. Alternatively, even if labor and intermediates are gross complements, control regions might lose market share if the cost reductions in Miami and Miami's customers steal business from firms in these control regions, again depressing control region labor demand.

These mechanisms reconcile the apparent contradiction between economic theory and empirical findings. Labor demand can indeed slope downward in Miami, as Borjas (2003) argues, while generating no detectable differences between treatment and control regions, as Card (1990); Peri and Yasenov (2019) observe. Unfortunately, the lack of comprehensive firm-level input-output data for the US prevents direct testing of these network-based explanations. Nonetheless, my results demonstrate that production network spillovers can fundamentally alter the interpretation of spatial variation designs in immigration research.

5.2 Does Where Immigrants Live Matter for Welfare?

Several host countries, including Canada, Germany, Sweden, Norway, and Finland, actively direct refugees and asylum seekers to specific regions, often to prevent overcrowding. My analysis of varying spillover effects raises a natural question: could there be meaningful welfare gains from concentrating immigrants in regions that are central to the production network?

To investigate this question, I simulate the arrival of 10,000 immigrants in each of Turkey's 81 provinces. For each simulation, I calculate the changes in prices across the economy and the aggregate welfare gains. The regional welfare change $d \ln Y_r$ is given by:

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

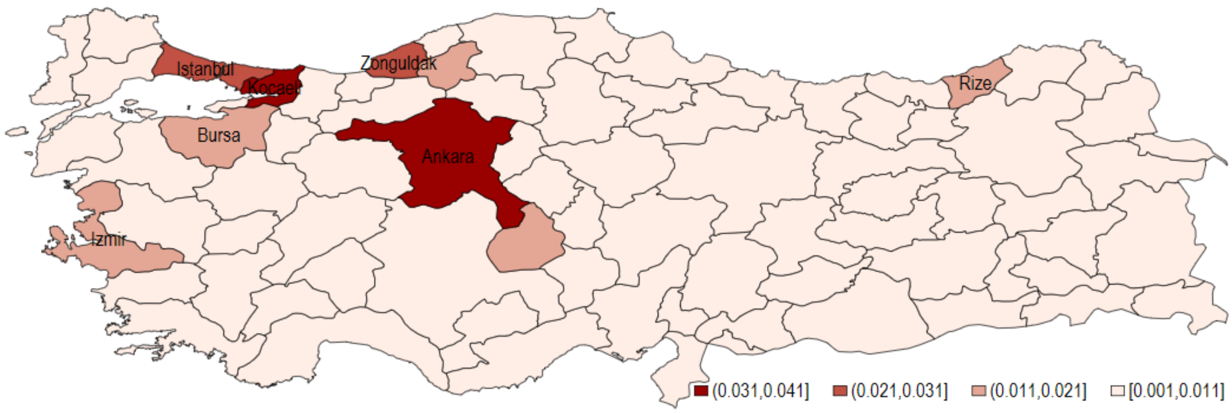
where welfare improves when either the region's share of total GDP (χ_r) increases or the prices of goods in its consumption basket decrease. I aggregate these regional welfare changes into a national measure using population-weighted averages.

Figure 7 maps the welfare effects of placing 10,000 immigrants (a 0.012% population increase for Turkey) across different provinces. While immigration increases welfare in all simulations, the magnitude varies dramatically by location. In most provinces, welfare gains are modest: less than 0.004%. However, when immigrants settle in Izmir, Zonguldak, Istanbul, Kocaeli, or Ankara, welfare gains range from 0.02% to 0.04%, more than ten times larger than typical effects. These provinces generate larger welfare gains because they are central nodes in the trade network. Their firms' extensive buying and selling relationships across regions allow immigration-induced cost reductions to benefit more regions, amplifying total welfare gains.

Notably, immigration to Kocaeli provides 50% larger welfare gains than the same number of immigrants going to Istanbul, despite Istanbul being more central. This occurs because Kocaeli's population is one-ninth of Istanbul's, so the same immigration shock lowers production costs in Kocaeli by a larger percentage, generating larger welfare gains in other cities. This highlights an important consideration: while the smallest, least-connected cities are suboptimal host locations from a welfare perspective, the largest cities are not necessarily optimal either. The optimal allocation depends on the precise structure of the production network.

These results should not be interpreted as advocating that all immigrants be directed to central network nodes. To maintain tractable exposition, I abstract from several important welfare considerations, including congestion externalities, housing supply elasticity, and local public goods provision. Future research should investigate the optimal spatial allocation of immigrants that incorporates these additional economic forces. My argument is more limited in its scope: current policy frameworks typically consider job availability in host regions and population balance for equity reasons. However, policymakers appear unaware that the centrality of the production net-

Figure 7: Heatmap of total welfare effects of immigration across host regions



Notes: This figure shows the results from 81 counterfactuals, one for each province of Turkey. Each counterfactual consists of an arrival of 10,000 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.

work amplifies immigration's economic benefits, enabling broader populations to take advantage of immigration-induced cost reductions. This represents a previously unrecognized first-order effect that should inform spatial allocation decisions. Understanding how the position of the network shapes the aggregate impact of immigration provides policymakers with a new lens for evaluating placement strategies, particularly when there are multiple feasible destinations.

5.3 Does the Skill Composition of Immigrants Impact Spillover Effects?

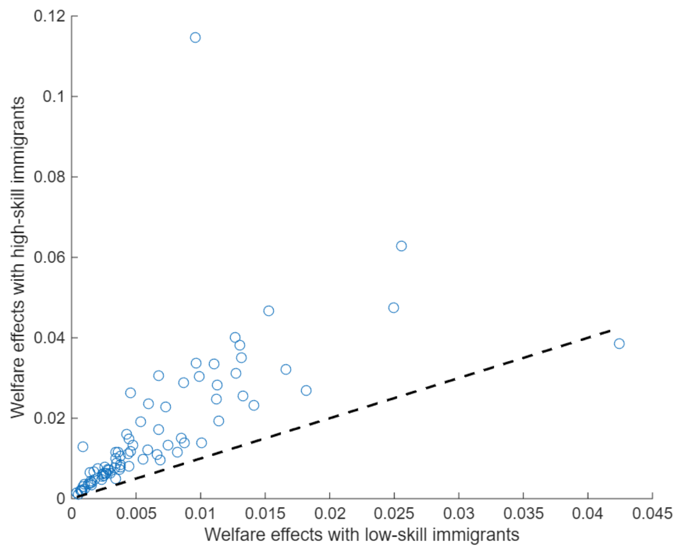
Immigration shocks often involve skill-specific labor supply changes. Syrian immigrants in Turkey, for example, have lower average education levels than natives and work in less skill-intensive industries like textiles, construction, and agriculture (Turkish Red Crescent and WFP, 2019). When low-skill and high-skill labor are imperfect substitutes, immigrants of different skill levels affect production costs in different industries. The magnitude of spillovers may therefore depend on the extent to which these affected industries trade with other regions.

To analyze skill-specific effects, I extend the baseline model to incorporate both low- and high-skill labor, with details provided in the Appendix Section A.3. One important caveat is that I must assume the substitution elasticity between low- and high-skilled workers ($\sigma_S = 1$) because the employer-employee matched data do not show the education of the workers.

To examine how spillovers vary with immigrant skill levels, I conduct paired counterfactuals for each of Turkey's 81 provinces. For each province, I simulate two scenarios: one with 10,000 low-skill immigrants and another with 10,000 high-skill immigrants, and then compare the resulting welfare effects.

Figure 8 compares the welfare effects of low-skill versus high-skill immigration. Each circle represents one of the 81 NUTS-3 regions, with low-skill immigration effects on the x-axis and high-skill effects on the y-axis. The dashed 45-degree line represents equal welfare effects; points above this line indicate regions where high-skill immigration generates larger welfare gains.

Figure 8: Comparison of welfare effects across low-skill and high-skill immigration



Notes: This figure shows the results from 162 counterfactuals, two for each NUTS-3 region in Turkey. For each region, I calculate the total welfare change when (1) 10,000 low-skill immigrants arrive in the host region and (2) 10,000 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.

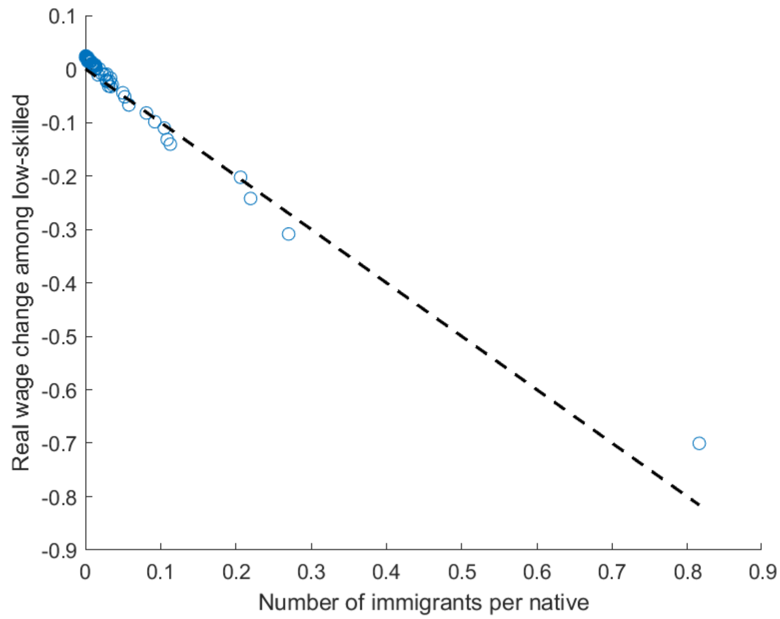
Figure 8 reveals two key patterns. First, most regions generate negligible welfare effects from immigration regardless of skill composition. These minimal effects reflect the limited spillover capacity of non-central regions: when host regions lack strong inter-regional trade connections, cost reductions fail to propagate through the production network, which constrains the welfare gains.

Second, central regions exhibit stark skill-based differentials in welfare generation. For example, a 10,000-immigrant shock in Bursa increases total welfare by 0.064% when immigrants are highly skilled versus 0.026% when low-skilled, a 2.5-fold difference. This differential stems directly from the structure of Turkey's production network: high-skill-intensive industries maintain stronger inter-regional trade linkages than their low-skill counterparts. Consequently, wage reductions in high-skill-intensive sectors propagate more extensively through the economy, amplifying the aggregate welfare impact of high-skill immigration to central nodes.

5.4 Model-based Factual: Quantifying the General Equilibrium Effects of Syrian Immigration

The counterfactuals suggest that immigration spillovers are largest when host regions are central in the trade network and immigrants are high-skilled. Since Syrian immigrants are concentrated in less-developed southeastern regions and have lower skill levels than native Turkish workers, one would expect limited general equilibrium effects.

Figure 9: Partial vs general equilibrium 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 A.3. Each blue circle denotes a Turkish province. The dashed line is the -45° line.

To test this prediction, I calculate how the low-skilled Syrian immigration affects Turkish natives' real wages and compare these general equilibrium effects with partial equilibrium predictions. Figure 9 plots this comparison across Turkey's 81 provinces, showing changes in low-skill natives' real wages (y-axis) against the 2019 Syrian-to-native ratio (x-axis). The dashed -45°line represents what would be observed with only partial equilibrium effects. The actual estimates closely track this line: the correlation between wage changes and immigration intensity is -0.99 (R-squared of 0.97), indicating that partial equilibrium effects accurately predict general equilibrium outcomes.

This finding validates prior studies of Syrian immigration's labor market effects in Turkey. Both Gulek (2024) and Gulek and Vives-i Bastida (2024) document the displacement of low-skill natives by Syrian immigrants. Their results accurately capture these effects because Syrian immigrants settled in regions non-central to Turkey's trade network, where SUTVA violations are minimal.

6 Conclusion

This paper formalizes how immigration-induced wage changes propagate through production networks, isolates the relevant economic forces, tests them empirically using novel methods, and quantifies immigration's total effects on wages and welfare across regions. The results help resolve persistent puzzles in the immigration literature, guide practitioners in identifying potential SUTVA violations in spatial designs, and inform policy decisions regarding immigrant allocation.

The theoretical model tractably demonstrates how cost reductions from immigration in one region can impact firms elsewhere. However, it purposefully abstracts from several established immigration effects that prior research has documented. Immigrants can affect skill-biased technological change (Lewis, 2011), boost productivity through entrepreneurship (Akcigit et al., 2017), alter product composition through different consumption patterns (Galaasen et al., 2025), influence offshoring decisions (Olney and Pozzoli, 2021), affect markups through enhanced price competition (Kim et al., 2025), and impact firms' monopsony power through differential bargaining positions (Naidu et al., 2016). In principle, all these forces can propagate through input-output networks, with network structure potentially amplifying their effects. Future work should investigate these mechanisms both theoretically and empirically.

My finding that large cities generate substantial spillovers has significant implications for practitioners. Immigration to smaller, less developed regions typically produces minimal spillovers, allowing traditional difference-in-differences analyses to capture local effects accurately. This explains why studies of refugee settlement in border regions—such as Syrians in southeastern Turkey (Gulek, 2024) or Venezuelans along the Colombian border (Bahar et al., 2024)—yield reliable results using spatial variation designs. Conversely, economic migration often targets larger, more connected cities. European hubs like Brussels, Frankfurt, and Munich, which maintain the highest foreign-to-native ratios in the EU (Mayors of Europe, 2019), likely generate significant spillovers throughout Europe, potentially biasing traditional empirical approaches that assume spatial independence.

Finally, while my model demonstrates that immigrants can reduce wages in host regions and harm some natives, it also reveals that immigration generates substantial welfare gains when immigrants settle in central regions. The former has dominated academic and policy discourse, while the latter represents a novel finding that can provide important nuance to contemporary immigration debates. This distinction proves particularly relevant today, as anti-immigrant sentiment has intensified globally (Hangartner et al., 2019; Alesina and Tabellini, 2024), making it crucial for policymakers to understand immigration's broader economic benefits alongside its distributional consequences.

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.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas, "Immigration and the rise of American ingenuity," *American Economic Review*, 2017, *107* (5), 327–331.

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.

Albert, Christoph and Joan Monras, "Immigration and spatial equilibrium: the role of expenditures in the country of origin," *American Economic Review*, 2022, *112* (11), 3763–3802.

Alesina, Alberto and Marco Tabellini, "The political effects of immigration: Culture or economics?," *Journal of Economic Literature*, 2024, *62* (1), 5–46.

Baez, Javier E, "Civil wars beyond their borders: The human capital and health consequences of hosting refugees," *Journal of Development Economics*, 2011, *96* (2), 391–408.

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://ahmetgulek.github.io/Bahar_DiTella_Gulek_Colombia.pdf.

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.

Beerli, Andreas, Jan Ruffner, Michael Siegenthaler, and Giovanni Peri, "The abolition of immigration restrictions and the performance of firms and workers: evidence from Switzerland," *American Economic Review*, 2021, *111* (3), 976–1012.

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.

Bonacich, Phillip, "Power and centrality: A family of measures," *American Journal of Sociology*, 1987, *92* (5), 1170–1182.

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.

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.

_ , "Immigration and inequality," *American Economic Review*, 2009, *99* (2), 1–21.

Carpio, Ximena V Del and Mathis Wagner, *The impact of Syrians refugees on the Turkish labor market*, The World Bank, 2015.

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.

Clemens, Michael A and Jennifer Hunt, "The labor market effects of refugee waves: reconciling conflicting results," *ILR Review*, 2019, *72* (4), 818–857.

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.

— , — , **and** — , "Labor supply shocks, native wages, and the adjustment of local employment," *The Quarterly Journal of Economics*, 2017, *132* (1), 435–483.

Edo, Anthony and Cem Özgüzel, "The impact of immigration on the employment dynamics of European regions," *Labour Economics*, 2023, *85*, 102433.

Galaasen, Sigurd, Andreas R Kostøl, Joan Monras, and Jonathan Vogel, "The labor supply curve is upward sloping: The effects of immigrant-induced demand shocks," 2025. NBER WP: https://www.nber.org/papers/w33930.

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://jvivesb.github.io/jmp/draft_siv.pdf.

Hangartner, Dominik, Elias Dinas, Moritz Marbach, Konstantinos Matakos, and Dimitrios Xefteris, "Does exposure to the refugee crisis make natives more hostile?," *American Political Science Review*, 2019, *113* (2), 442–455.

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

Jackson, Matthew O, *Social and economic networks*, Vol. 3, USA: Princeton University Press, 2008.

Jaeger, David A, Joakim Ruist, and Jan Stuhler, "Shift-share instruments and the impact of immigration," 2018. NBER WP: https://www.nber.org/papers/w24285.

Kim, Ryan, Justin H Leung, and Ariel Weinberger, "Prices and immigration: A firm-Level analysis," 2025. Available at: https://www.freit.org/RMET/2025/selected/weinberger.pdf.

Lewis, Ethan, "Immigration, skill mix, and capital skill complementarity," *The Quarterly Journal of Economics*, 2011, *126* (2), 1029–1069.

_ **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.

Mayors of Europe, "Top 10 cities with the highest share of foreigners in population," https://mayorsofeurope.eu/reports-analyses/eu-cities-with-most-immigrants/ 2019. Accessed: (1/4/2025).

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

Naidu, Suresh, Yaw Nyarko, and Shing-Yi Wang, "Monopsony power in migrant labor markets: evidence from the United Arab Emirates," *Journal of Political Economy*, 2016, *124* (6), 1735–1792.

Olney, William W and Dario Pozzoli, "The impact of immigration on firm-level offshoring," *Review of Economics and Statistics*, 2021, *103* (1), 177–195.

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," *Review of Economics and Statistics*, 2025, pp. 1–29.

Turkish Red Crescent and WFP, "Refugees In Turkey: Livelihoods survey findings," 2019. https://reliefweb.int/report/turkey/refugees-turkey-livelihoods-survey-findings-2019-entr.

Ulyssea, Gabriel, "Firms, informality, and development: Theory and evidence from Brazil," *American Economic Review*, 2018, *108* (8), 2015–47.

_ , "Informality: Causes and consequences for development," *Annual Review of Economics*, 2020, *12*, 525–546.

A Model Appendix

A.1 Proofs

Before showing the proofs, I introduce some notation. The trade matrix Ω is of size $(N+R) \times (N+R)$, where the last R rows are zeros. I 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, I 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, I 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, I 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, I get:

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

■

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, I 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, I get:

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

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, I can write

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

which implies

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

Focusing on the first part of equation 16, I can write χ_{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, I can show:

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

Hence, I need to calculate $d\Omega$. First, using CES algebra, I 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, I 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, I know:

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

More succinctly, I 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 $d\Omega_{i,j}$, I 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 $d\Psi = \Psi d\Omega \Psi$, I 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, I 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.

■

A.2 Counterfactuals Without Demand Effects

In the main model, immigrants supply labor and consume goods. It is of general interest to separate the effects of these two actions. To calculate the effects of a scenario in which immigrants supply labor but do not consume goods, I change the system of linear equations in a small way.

$$
\begin{aligned}
d\ln\lambda_r =& \sum_{i\in r}\frac{\lambda_i\Omega_{iL}}{\lambda_r}(d\ln\lambda_i + d\ln\Omega_{iL}),\\
d\ln\Omega_{i,L} =& (1-\sigma_u)(d\ln w_{r_i} - \sum_{j=1}^{n}\frac{\tilde{\Omega}_{ij}}{\tilde{\Omega}_{iM}}d\ln p_j),\\
d\ln\lambda_i =& (1-\sigma_l)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i\mu_j}Cov_{\tilde{\Omega}^{(j)}}\left(d\ln p, \Psi_{(i)}\right),\\
&+(\sigma_u-\sigma_l)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\tilde{\Omega}_{j,L}\left(d\ln w_{r_j} - \sum_{k=1}^{n}\frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}}d\ln p_k\right)(\Psi_{ji}-I_{ji}),\\
&+\frac{1}{\lambda_i}\sum_j\sum_r b_{rj}\Psi_{ji}\chi_r d\ln\chi_r,\\
d\ln\chi_r =& \left(\sum_{i\in r}\frac{\pi_i}{\chi_r}d\ln\lambda_i\right)+\frac{\lambda_r}{\chi_r}d\ln w_r,\\
d\ln p_i =& \sum_{j=1}^{n}\tilde{\Psi}_{ij}\tilde{\Omega}_{jL}d\ln w_{r_j},\\
d\ln w_r =& d\ln\lambda_r - d\ln L_r.
\end{aligned}
\tag{17}
$$

where on line 4, I change the change in regional labor income $d\ln\lambda_r$, which equals the change in wages and labor, into just the change in wages $d\ln w_r$.

A.3 Model with Skill Heterogeneity

A.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 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. I 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 σ_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. I assume that firms have constant and exogenous markup μ_i, and therefore set price $p_i = \mu_i C_i$.

Note that this specification extends the baseline model by introducing skill heterogeneity through an additional CES nest for labor types. To maintain notational clarity, I adopt the following convention: σ_u continues to denote the elasticity of substitution between labor and intermediates in the upper nest, while $\sigma_m \equiv \sigma_l$ denotes the elasticity between intermediates in the lower nest (relabeled from σ_l to avoid confusion with the new parameter σ_L, which governs substitution between low- and high-skill labor). This three-tier structure allows the model to capture both skill-biased effects of immigration and their differential propagation through the production network.

Final Demand

All final goods consumption as well as the ownership of firms is local. I 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.[19]

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.

A.3.2 Input-Output Definitions

I 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

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

[19] Note that one could also model the consumption of low- and high-skill workers separately, which would me to track welfare effects on these worker types separately. However, since the focus is on the wage effects, I abstract away from this detail.

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*

$$d\ln\tilde{\Omega}_{i,L} = (1-\sigma_u)\left(\frac{1-\tilde{\Omega}_{i,L}}{\tilde{\Omega}_{i,L}}(\tilde{\Omega}_{il}d\ln w_{rl} + \tilde{\Omega}_{ih}d\ln w_{rh}) - \sum_{j=1}^{n}\tilde{\Omega}_{ij}d\ln p_j\right) \tag{18}$$

Proof follows the same steps as in the baseline model.

Let $\overline{w_j} = \frac{1}{\tilde{\Omega}_{j,L}}\left(\tilde{\Omega}_{j,l}d\ln w_{r_j,l} + \tilde{\Omega}_{j,h}d\ln w_{r_j,h}\right)$ denote 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,*

$$\begin{aligned} d\ln\lambda_i =& \sum_{j=1}^{n}(1-\sigma_M)\frac{\lambda_j}{\lambda_i\mu_j}Cov_{\tilde{\Omega}^{(j)}}\left(d\ln p, \Psi_{(i)}\right) \\ &+ (\sigma_U-\sigma_M)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\tilde{\Omega}_{j,l}\left(d\ln\overline{w_{r_j}} - \sum_{k=1}^{n}\frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}}d\ln p_k\right)(\Psi_{ji}-I_{ji}) \\ &+ \frac{1}{\lambda_i}\sum_j\sum_r\overline{b_{rj}}\Psi_{ji}\chi_r d\ln\chi_r, \end{aligned} \tag{19}$$

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,*

$$d\ln p_i = \sum_{j=1}^{n}\tilde{\Psi}^p_{i,j}(\tilde{\Omega}_{j,l}w_{r_j,l} + \tilde{\Omega}_{jh}w_{r_j,h}) = \sum_{f=1}^{F}\tilde{\Psi}_{i,f}d\ln w_f, \tag{20}$$

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, I 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.*

$$
\begin{aligned}
d\ln w_f =& d\ln\lambda_f - d\ln L_f \\
d\ln p_i =& \sum_{f=1}^{F} \tilde{\Psi}_{i,f} d\ln w_f \\
d\ln\lambda_i =& \sum_{j=1}^{n} (1-\sigma_M)\frac{\lambda_j}{\lambda_i\mu_j} Cov_{\tilde{\Omega}^{(j)}}\left(d\ln p, \Psi_{(i)}\right) \\
&+ (\sigma_U - \sigma_M)\sum_{j=1}^{n}\frac{\lambda_j}{\lambda_i}\tilde{\Omega}_{j,l}\left(d\ln\overline{w_{r_j}} - \sum_{k=1}^{n}\frac{\tilde{\Omega}_{j,k}}{\tilde{\Omega}_{j,M}} d\ln p_k\right)(\Psi_{ji} - I_{ji}) \\
&+ \frac{1}{\lambda_i}\sum_j\sum_r \overline{b_{rj}}\Psi_{ji}\chi_r d\ln\chi_r \\
d\ln\chi_r =& \left(\sum_{i\in r}\frac{\pi_i}{\chi_r} d\ln\lambda_i\right) + \frac{\lambda_r}{\chi_r} d\ln\lambda_r \\
d\ln\lambda_r =& s^w_{ls} d\ln\lambda_{rl} + s^w_{hs} d\ln\lambda_{rh} \\
d\ln\lambda_f =& \sum_{i=1}^{n}\frac{\lambda_i\Omega_{if}}{\lambda_f} d\ln\lambda_i + \sum_{i=1}^{n}\frac{\lambda_i\Omega_{if}}{\lambda_f} d\ln\Omega_{if} \\
d\ln\tilde{\Omega}_{il} =& d\ln\tilde{\Omega}_{iL} + (1-\sigma_L)\left[d\ln w_{r_i,l} - \frac{1}{\tilde{\Omega}_{iL}}\left(\tilde{\Omega}_{il} d\ln w_{r_i,l} + \tilde{\Omega}_{ih} d\ln w_{r_i,h}\right)\right] \\
d\ln\tilde{\Omega}_{ih} =& d\ln\tilde{\Omega}_{iL} + (1-\sigma_L)\left[d\ln w_{r_i,h} - \frac{1}{\tilde{\Omega}_{iL}}\left(\tilde{\Omega}_{il} d\ln w_{r_i,l} + \tilde{\Omega}_{ih} d\ln w_{r_i,h}\right)\right] \\
d\ln\tilde{\Omega}_{iL} =& (1-\sigma_u)\left[\frac{1-\tilde{\Omega}_{iL}}{\tilde{\Omega}_{iL}}(\tilde{\Omega}_{il} d\ln w_{rl} + \tilde{\Omega}_{ih} d\ln w_{rh}) - \sum_{j=1}^{n}\tilde{\Omega}_{ij} d\ln p_j\right]
\end{aligned} \tag{21}
$$

Proof follows the same steps as for the baseline model.

B Data Appendix

B.1 Calculating Trade Exposures

This section describes the calculation of firm-level and region-industry level trade exposures.

For computational tractability in the counterfactual analyses, I work with representative firms at the region-industry level. I define regions as Turkey's 81 provinces and 86 industries at the two-digit classification level. Given that not every industry operates in each province, I obtain 4185 region-industry cells in total. Computing the trade exposures $[U, D1, D2]$ requires inverting 4185×4185 matrices, which remains computationally feasible at this level of aggregation.

Working with firm-level data introduces several considerations that merit detailed discussion.

B.1.1 Data Cleaning

Trade relationships, transaction volumes, sales, and labor payments exhibit substantial year-to-year variation. To minimize measurement error, I focus on firms that maintain continuous operations and file complete balance sheet statements throughout the pre-period from 2006 to 2010. I further restrict the sample to firms reporting positive labor costs in the tax records, excluding firms without employees. This yields 285,178 firms in the analysis sample, some with observed trading relationships and others without.

I define material costs as total firm purchases recorded in the VAT data, encompassing both domestic purchases and imports. Firms making no intermediate purchases receive a labor share equal to one.

The data reveal 9006 firms (3.2% of the sample) whose total costs (intermediate purchases plus labor costs) exceed reported sales, even when aggregated over the entire 2006–2010 pre-period. This pattern poses challenges for the model's validity. Well-defined Leontief inverse matrices require that each row sum in the trade matrices remains less than or equal to one, a condition violated when costs exceed sales. I address this issue by treating sales information from balance sheets as measured with error relative to the more precisely recorded labor costs and purchases from tax records. Accordingly, I adjust these firms' sales upward to equal their total costs, effectively imposing a markup of one.

B.1.2 Measuring Upstream Exposure

The upstream exposure effect on labor share follows from the model:

$$U_i = \sum_{k=1}^{n} \tilde{\Psi}_{ik} \tilde{\Omega}_{k,L} \delta_{r_k}$$

Implementing this measure using the full firm-level network would require inverting a $285,000 \times 285,000$ matrix. While the cost-based trade matrix $\tilde{\Omega}$ is sparse, its Leontief inverse $\tilde{\Psi}$ is not. Moreover, the firm-level network approach cannot capture new trading relationships that may form endogenously in response to the immigration shock. Region-industry level trade matrices prove more robust to such endogenous network formation than firm-to-firm matrices. As Appendix Section C.1 demonstrates, most new trading relationships form within rather than across regions. This implies that even when cross-regional relationships do emerge, they occur predominantly between firms that already maintain trading connections with that region.

However, using region-industry level aggregation for the reduced-form analysis reduces statistical power and may compromise identification credibility. Consider comparing textile and finance firms in Istanbul because textile firms typically purchase from southeast regions while finance firms do not. This comparison provides less compelling identification than comparing two textile firms in Istanbul that happen to differ in their exposure to immigrant-receiving regions through their specific trading partners.

To balance these competing considerations, I calculate firm-level trade exposures while exploiting region-industry level variation in trading patterns. Two textile firms in Istanbul exhibit differential upstream exposure if one sources inputs from host regions in the southeast while the other sources from non-host regions. For firms purchasing identical quantities from a particular region-industry cell, I do not distinguish between those buying from labor-intensive versus non-labor-intensive suppliers, as changing trading partners across labor intensity dimensions proves more feasible than changing regional sourcing patterns.

Formally, I calculate upstream exposure as:

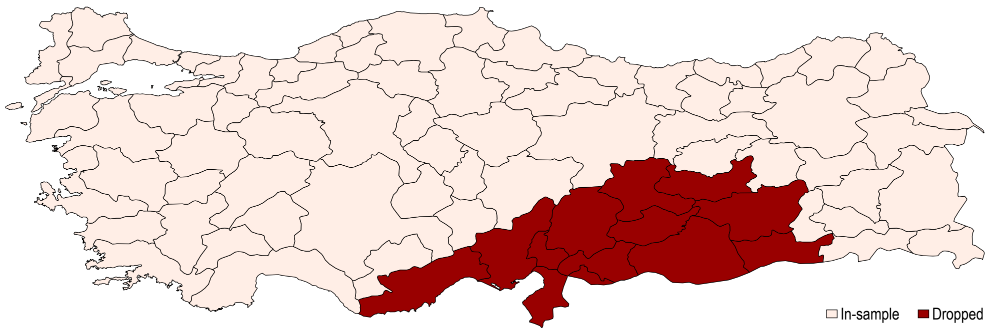
$$U_{it} = \sum_{j=1}^{J} \tilde{\Psi}_{Nj} \times \tilde{\Omega}_{j,L} \times \delta_{jt}, \tag{22}$$

where $\tilde{\Psi}_{NJ}$ represents the firm-by-region-industry trade matrix whose ijth element captures firm i's cost dependence on region-industry j, $\tilde{\Omega}_{j,L}$ denotes the labor share of region-industry j, and δ_{jt} measures the region-industry-time level immigration shock.

This approach offers several advantages. First, the $\tilde{\Psi}_{NJ}$ matrix requires substantially less computational resources than $\tilde{\Psi}_{NN}$. Second, while upstream exposure retains firm-level variation, it primarily exploits region-industry level differences in trading relationships.

Having defined upstream exposure using region-industry level trading partners, the downstream exposures follow naturally. Both $D1$ and $D2$ depend on how firms' prices and costs respond to the shock, captured by $d \ln p$. I calculate $d \ln p$ using the region-industry level framework and then construct the downstream exposure measures accordingly.

Figure B.1: Regions omitted from the main analysis



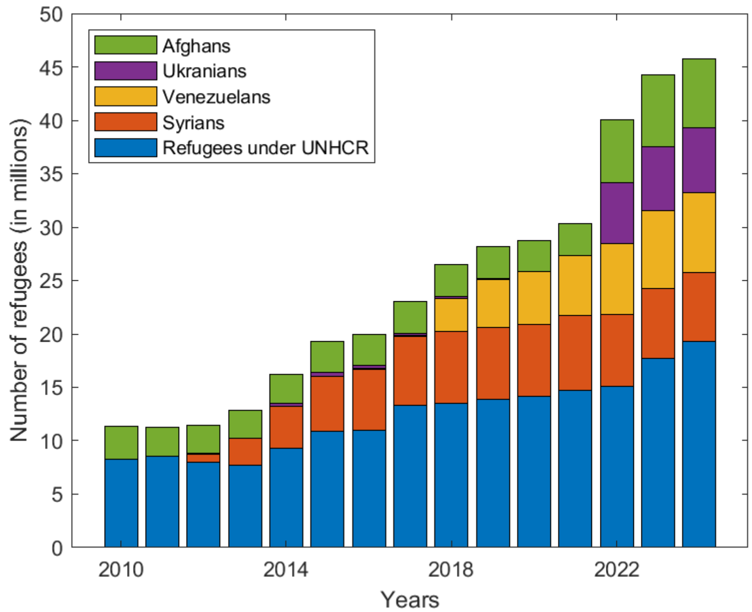
B.2 Summary Statistics

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 50 employees in 2010				
217.74	4.00	68	0.71	0.16
(495.77)	(14.51)	(419)	(0.45)	(0.15)

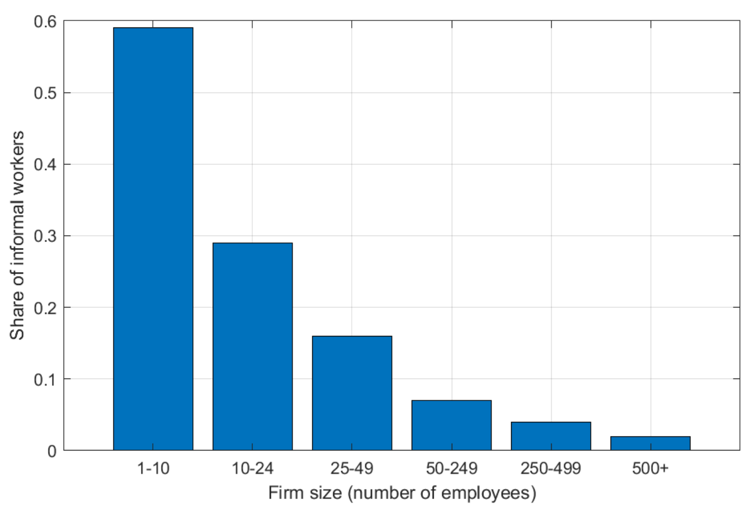
Note: Data is resticted to Manufacturing firms in non-exposed regions that exist throughout 2006–2019. There are 19505 such firms in the sample. 1112 of these firms have more than 50 employees in 2010. Standard deviations are in parenthesis.

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



Source: Author's calculations using UNHCR data. This dataset is publicly available from https://www.unhcr.org/refugee-statistics

Figure B.3: The Ratio of Informal Workers Across Firm Size



Source: Author's calculations using HLFS. Informality is defined as the ratio of workers who self-declare that their employer does not pay for social security, which is legally mandated for all formal workers in Turkey.

C Supporting Evidence

C.1 New Trade Formation

The model assumes constant trade networks. While firms can adjust trade volumes with existing partners, the model does not incorporate endogenous network formation. This simplification is unlikely to affect my results because regional connections matter more than specific firm identities: firms predominantly form new supplier relationships within regions where they already have trade partners. I provide both descriptive and regression evidence for this claim.

Descriptive Evidence

Turkey's trade network exhibits strong regional persistence. Of the 6.4 million buyer-seller connections in 2010, 4.0 million (62%) were new since 2008. These new connections display two striking patterns. First, 53% occurred within regions, reflecting the dominance of local trade. Second, among cross-regional connections, 59% involved buyers who already had suppliers in the destination region. Thus, even when firms expand beyond their home region, they overwhelmingly choose regions where they maintain existing relationships.

Regression Analysis

To control for firm characteristics and regional factors, I estimate how pre-existing regional connections predict future network formation. Using all firms in non-host regions, I identify whether each firm had a supplier from a host region in 2010. I then track whether these firms established new supplier relationships by 2019, distinguishing between new connections in host regions (primary outcome) and non-host regions (placebo outcome). I estimate:

$$y_{irj} = \beta D_i + \theta W_{irj} + \epsilon_{irj}, \tag{23}$$

where D indicates having a pre-existing supplier in a host region in 2010, and W includes cubic polynomials of supplier count and region-by-industry fixed effects. This specification compares firms within the same region-industry cell, which isolates the predictive power of existing regional connections.

Table C.2 presents the results. Panel A shows that firms with pre-existing host-region suppliers are 22–40 percentage points more likely to form new relationships in host regions, with larger effects for manufacturing firms. These estimates remain stable across specifications and are statistically significant at the 5% level throughout.

Panel B provides a placebo test using new connections in non-host regions as the outcome. If pre-existing connections merely proxied for firms' general propensity to form new relationships, we would observe similar positive coefficients. Instead, all estimates are small, negative, and statistically insignificant, confirming that the effect is specific to host regions.

Table C.2: The Effects of Having an Existing Regional Supplier on New Trade Formation

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: New supplier in host region						
Pre-existing regional supplier	0.279***	0.214***	0.224***	0.376*	0.387*	0.398*
	(0.052)	(0.048)	(0.047)	(0.187)	(0.185)	(0.179)
Panel B: New supplier in non-host region						
Pre-existing regional supplier	-0.029	-0.062	-0.021	-0.046	-0.042	-0.033
	(0.042)	(0.041)	(0.042)	(0.132)	(0.128)	(0.133)
Sample	All	All	All	Man	Man	Man
N	97484	97487	96747	20172	20172	19944
Controls						
Cubic of number of suppliers	N	Y	Y	N	Y	Y
Region-by-industry F.E.	N	N	Y	N	N	Y

Notes: In Panel A, the outcome is an indicator of whether the firm established a new supplier relationship in 2019 with a firm from a host region, where "new" is defined as a relationship that did not exist in 2008-2009. In Panel B, the outcome is an indicator of whether the firm established a new supplier relationship with a firm in a non-host region in 2019. The treatment variable is an indicator for the buyer firm having at least one supplier in host regions in 2010. The sample includes all firms in columns 1–3 and is restricted to manufacturing firms in columns 4–6. Robust standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001

The evidence demonstrates that regional connections, rather than specific firm relationships, drive network expansion. Firms with existing ties to host regions are substantially more likely to form additional connections there, validating my focus on region-industry-level rather than firm-level exposure measures.

C.2 Industry Heterogeneity in Structural Parameters

In the main text, I argue that structural elasticity estimates are common across industries. Here, I provide empirical evidence for this assumption. I approach this in two ways. First, I focus on large firms as in the main text, estimate equations 8 and 9 separately for each two-digit manufacturing industry, calculate elasticity parameters using GMM estimation, and apply Empirical Bayes shrinkage to avoid bias from small samples. This analysis results in relatively large confidence intervals due to the limited sample size of large firms in certain industries. As a secondary approach, I utilize the fact that the effect of upstream exposure on labor demand is similar among small and large firms, as shown in the Online Appendix Section D.5. Using this observation, I estimate equation 8 using all manufacturing firms to obtain estimates of the elasticity of substitution between labor and intermediates, which provides more precise estimates due to increased sample size.

Empirical Bayes

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

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

I 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, I obtain the Empirical Bayes posterior mean and variance:

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

Results

Figure C.4 presents the heterogeneity of elasticity of substitution estimates across manufacturing industries. Panel A shows the elasticity of substitution between labor and intermediates across two-digit manufacturing industries using large firms only. Panel B shows the elasticity of substitution between different intermediates using the same sample. Panel C presents the elasticity of substitution between labor and intermediates estimated using all manufacturing firms.

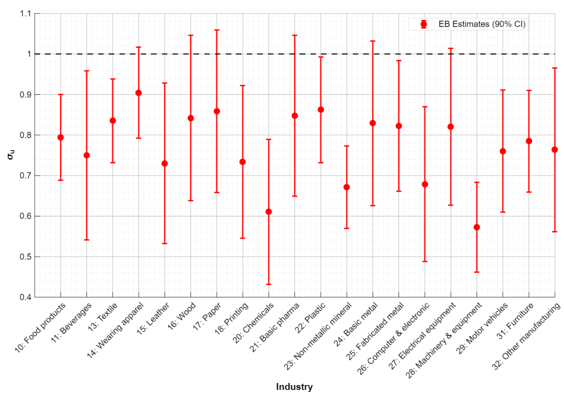
Labor-Intermediate Substitution (Panel A): The Empirical Bayes estimates show that labor and intermediates are gross complements across all manufacturing industries, with elasticities ranging from approximately 0.65 to 0.90. Most industries cluster around the aggregate estimate of 0.79 from the main text. The confidence intervals overlap substantially across industries, suggesting that the assumption of common elasticity parameters is reasonable for this technology parameter.

Intermediate-Intermediate Substitution (Panel B): The elasticity of substitution between different intermediates shows more heterogeneity across industries. While most industries have elasticities close to 1 (consistent with Cobb-Douglas production), some industries like textiles and basic metals exhibit elasticities above 1 (gross substitutes), while others like chemicals show elasticities below 1 (gross complements).

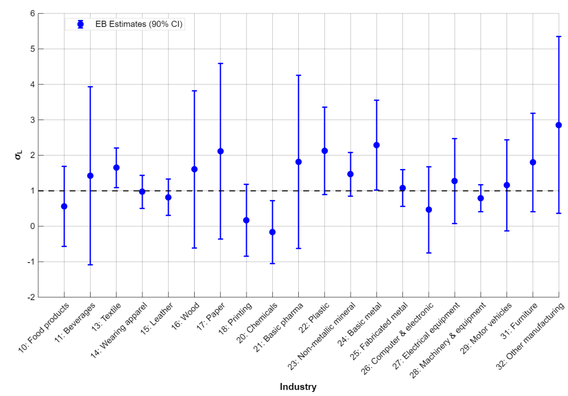
Robustness Using All Firms (Panel C): Panel C presents estimates using all manufacturing firms rather than just large firms, exploiting the finding that upstream exposure effects on labor demand are similar across firm sizes. The pattern mirrors Panel A: labor and intermediates are gross complements across all industries, with elasticities ranging from 0.70 to 0.85. The confidence intervals are narrower due to the larger sample size, and the cross-industry variation is smaller, providing additional support for the common elasticity assumption.

The results support the modeling assumption of common structural parameters across industries, particularly for the labor-intermediate substitution elasticity. While some heterogeneity exists, especially for intermediate-intermediate substitution, the estimates are generally consistent with the aggregate parameters used in the main analysis.

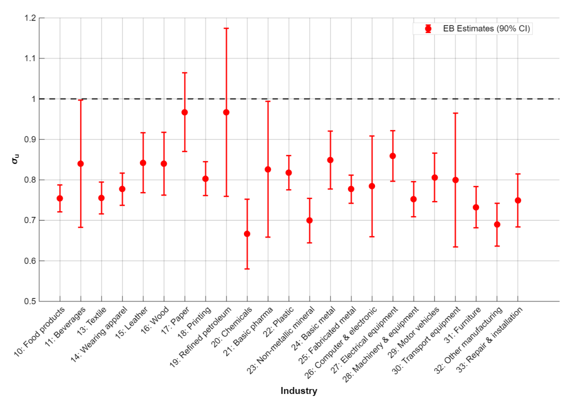
Figure C.4: Heterogeneity of EoS estimates across Manufucturing industries



(a) EoS between labor and intermediates



(b) EoS between intermediates



(c) EoS between labor and intermediates (all manufacturing firms)

Notes: Panels A and B report estimates from the same identification strategy as in the main text, which restricts the sample to manufacturing firms with 50+ employees in 2010, and estimates EoS parameters using GMM. Panel C reports estimates from an alternative design, which uses the sample from all manufacturing firms, and calculates EoS between labor and intermediates using the mapping in equation 10. Elasticity estimates using Empirical Bayes are reported. 90% confidence intervals are plotted.

C.3 Native Migration Responses

This section examines whether native population movements respond to Syrian immigration, a potential violation of my identification strategy. I estimate both event-study and IV specifications at the province level to test for native migration responses.

The event study specification is as follows:

$$y_{rt} = \sum_{t' \neq 2010} \beta_{t'} Z_r \mathbb{1}\{t = t'\} + f_r + f_t + \epsilon_{rt}. \tag{24}$$

The IV specification is as follows:

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

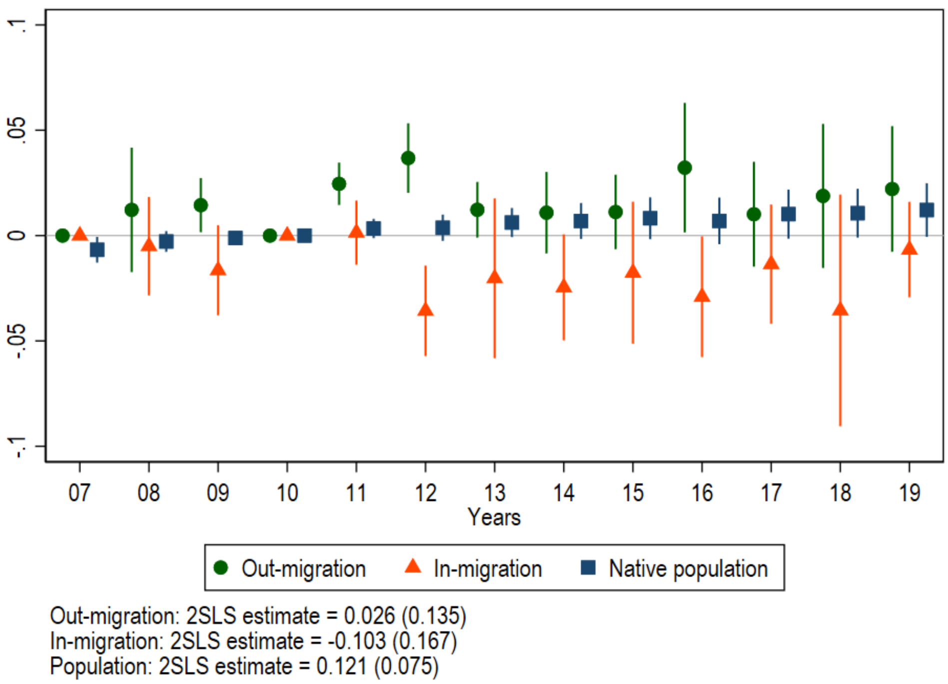
where y_{rt} represents the natural logarithm of in-migration, out-migration, or population; D_{rt} is the immigration treatment; Z_{rt} is the distance instrument; and f_r, f_t are region and time fixed effects.

Figure C.5 presents the results. The event-study estimates reveal modest migration responses concentrated in 2011–2012, before large-scale Syrian arrivals. In-migration rates decline slightly while out-migration rates increase marginally during these early years. By 2013 and beyond, when the Syrian presence becomes substantial, native migration responses are statistically indistinguishable from zero.

The economic magnitudes confirm that native migration cannot explain my results. A one standard deviation increase in the instrument, which corresponds to a 9% increase in the immigrant/native ratio by 2018, reduces in-migration by only 4%. Since annual in-migration constitutes less than 3% of local population in host regions, this translates to a negligible 0.01% population change per percentage point of immigration. Out-migration effects are similarly minimal.

These negligible migration responses allow southeastern regions to maintain their historical population growth trajectories. Despite hosting millions of Syrian refugees, these regions continue experiencing higher native population growth driven by persistently higher birth rates. Native labor mobility therefore plays no meaningful role in equilibrating regional labor markets following the immigration shock.

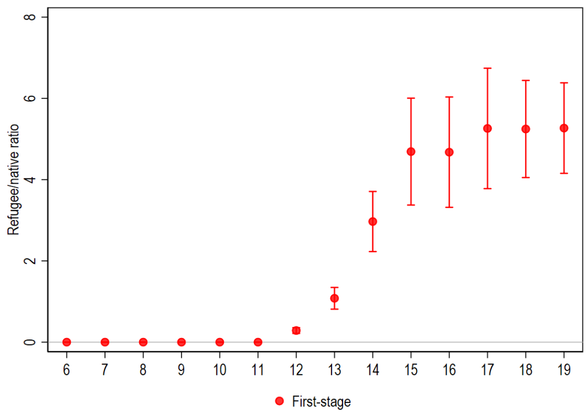
Figure C.5: Native migration responses to Syrian immigration



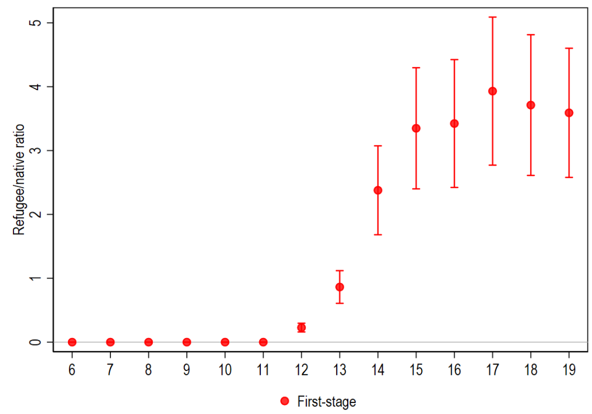
Notes: Event-study estimates from $y_{rt} = \sum_{t' \neq 2010} \beta_{t'} Z_r \mathbb{1}\{t = t'\} + f_r + f_t + \epsilon_{rt}$, where Z_r is the normalized distance share (standard deviation = 1) and f_r, f_t are region and time fixed effects. Outcomes are natural logarithms of in-migration, out-migration, and native population. Address-based tracking data begins in 2007. Therefore, estimates for native population start from 2007, and estimates from migration patterns start from 2008. 2010 is normalized because 2011 is the beginning of the Syrian Civil War. Standard errors are clustered at the region level. 95% confidence intervals are plotted.

C.4 Alternative Instruments

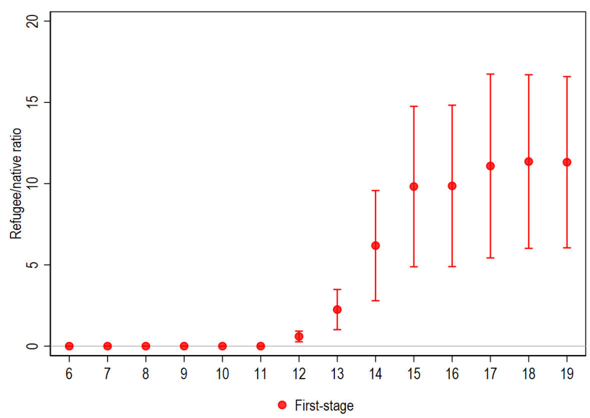
Figure C.6: First-stage estimates across different instruments



(a) First-stage: Distance



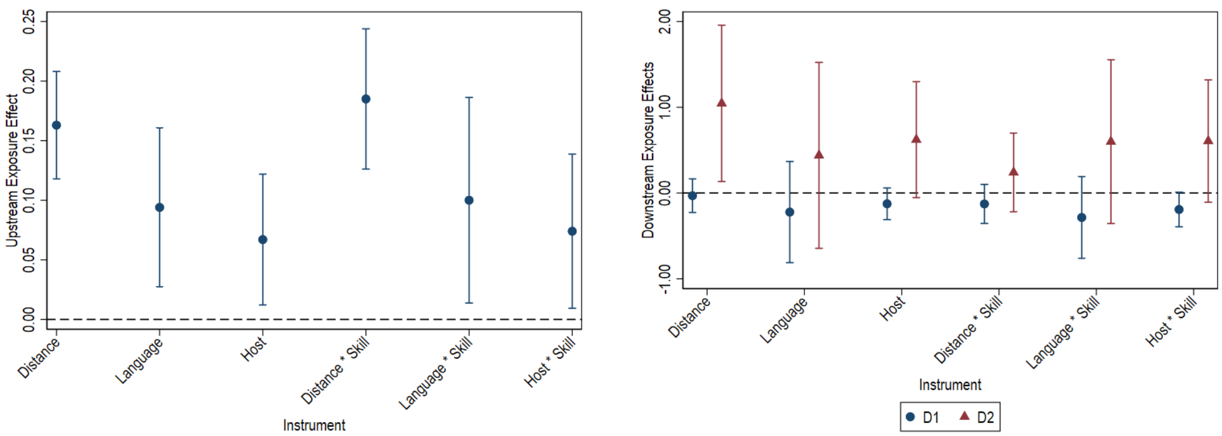
(b) First-stage: Language



(c) First-stage: Host

Source: Author's calculations. Distance and Language instruments are standardized to have standard deviation of one, while "host" instrument is simply a dummy indicator for the omitted regions. Figures show 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 I weight each region by its population in 2011. Standard errors are clustered at the region level. 95% confidence intervals are plotted.

Figure C.7: Effects of trade exposures on labor demand and sales under different instruments



(a) Labor share

(b) Sales

Notes: This figure shows the SIV estimates of the effects of upstream exposure on labor share and the effects of downstream exposures (D1 and D2) on sales, using different instruments. The regression equations are provided in equations 8 and 9 in the main text. Columns 1–3 use distance, language, and dummy shares in the construction of the shift-share instruments. In columns 4–6, I interact the regional shares with industry-level skill intensity, where skill intensity is measured as the share of workers without high school degrees in each industry. In these specifications, a firm has stronger upstream exposure if it purchases from firms operating in low-skill industries located in immigrant-hosting regions. Downstream exposures are calculated similarly. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

D Synthetic IV Robustness Checks

D.1 Comparisons between IV and SIV

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

I define the event-study equations for the IV estimator as follows: for labor share:

$$log(y_{isrt}^{L}) = \sum_{t' \neq 2010} \beta_{1,t'} U_i^Z \mathbb{1}\{t = t'\} + f_i^L + f_{srt}^L + W_{it}^L + \nu_{it}^L, \tag{25}$$

and for sales,

$$log(y_{isrt}^{S}) = \sum_{t' \neq 2010} \left(\beta_{t'}^{D1} D1_i^Z + \beta_{t'}^{D2} D2_i^Z\right) \mathbb{1}\{t = t'\} + f_i^S + f_{srt}^S + W_{it}^S + \nu_{it}^S. \tag{26}$$

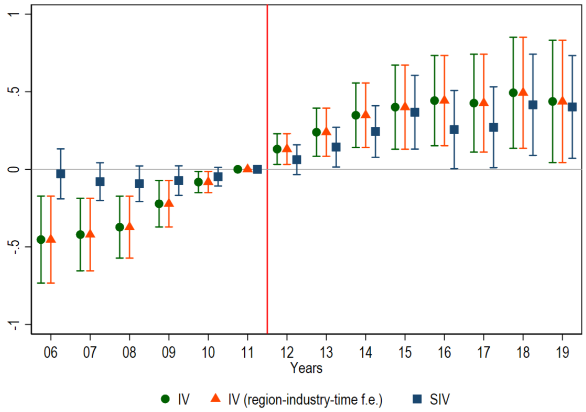
Here, f_{srt} denotes industry-region-time fixed effects that partial out industry-region level shocks. In robustness checks, I further group firms into quartiles based on their baseline sizes and control for industry-region-size-time fixed effects. This specification allows me to compare firms within the same region-industry cell that have similar numbers of employees at baseline.

Figure D.8 shows the results. In each panel, I plot estimates from the IV design with and without size-region-industry-time fixed effects, together with the baseline SIV estimates using distance-to-border as an instrument. Panels A, B, and C show upstream exposure effects on employment, wage bill, and labor share, respectively, while Panels D and E show downstream exposure effects on sales.

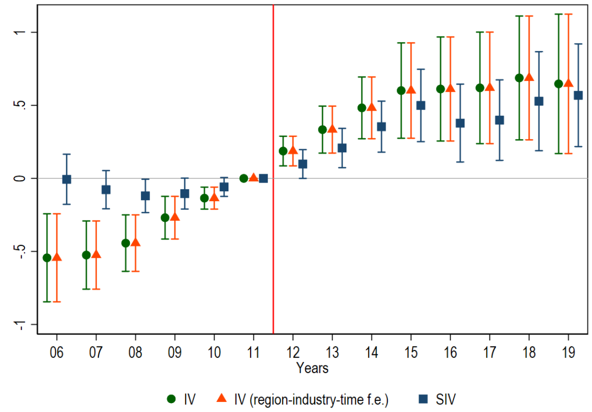
The results reveal that IV is biased in most, if not all, specifications. For example, more upstream-exposed firms followed differential trends from 2006 to 2011 in employment, wage bill, and labor share compared to less-exposed firms. This pattern persists even when controlling for region-industry-size-time fixed effects. Panel D also reveals significant pre-trends in the reduced-form analysis with baseline IV, where D1-exposed firms' sales grew differentially between 2006–2011 compared to less-exposed firms. These persistent differential trends motivate the use of Synthetic IV in the main text.

Panel E shows no pre-trends: more and less D2-exposed firms follow parallel trajectories before the immigration shock. However, this finding alone neither supports nor opposes the use of SIV. To establish causality, I require exogenous variation in all three treatment variables. Since pre-trends appear in two of the three cases, I cannot rely on IV for credible causal inference. This methodological challenge motivates the use of SIV.

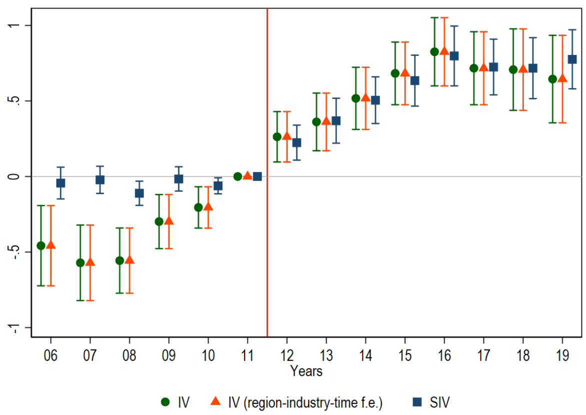
Figure D.8: Effects of trade exposures on firms' labor demand and sales



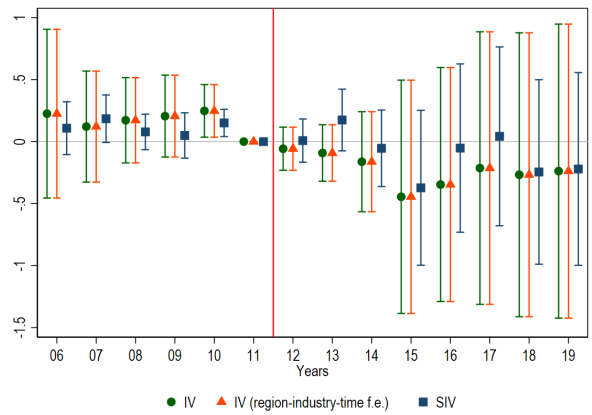
(a) Number of employees



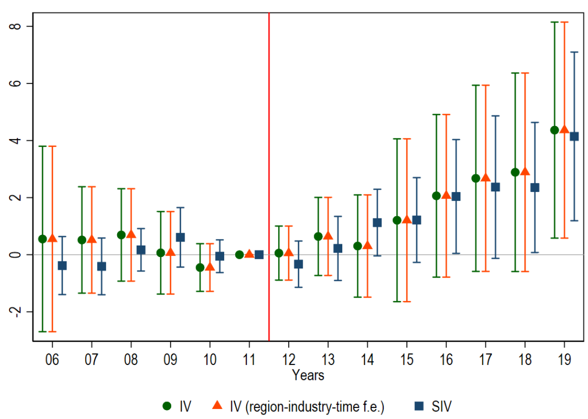
(b) Payroll



(c) Labor share



(d) Sales: Downstream-1



(e) Sales: Downstream-2

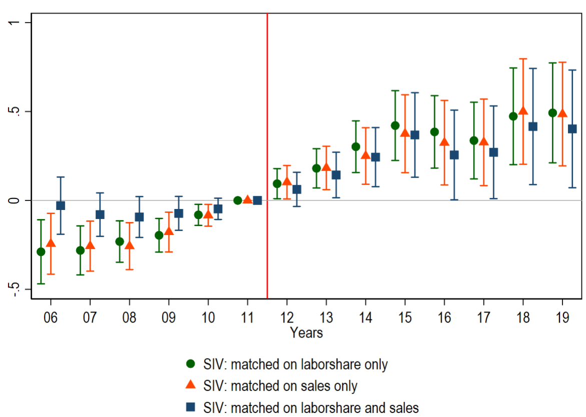
Notes: In Panels A, B, and C, the estimates come from the regression equation $y_{it} = \sum_{t' \neq 2011} \gamma_{1,t'} U_i^Z \mathbb{1}\{t = t'\} + f_i + f_t + \nu_{it}$, where the outcome variable is the natural logarithm of the number of workers in Panel A, of total payroll in Panel B, and of labor share in Panel C. In Panels D and E, the estimates come from the regression equation: $y_{it} = \sum_{t' \neq 2011} \left(\gamma_{2,t'} D1_i^Z + \gamma_{3,t'} D2_i^Z\right) \mathbb{1}\{t = t'\} + f_i + f_t + \nu_{it}$. SIV estimates use the debiased versions of the outcome and exposures, while IV uses the raw versions. SIV matching is based on the demeaned versions of labor share and sales. The instrument is based on the travel distance. Sample is restricted to manufacturing firms with at least 50 employees in 2010. Standard errors are clustered at the firm level. 95% confidence intervals are plotted.

D.2 Robustness to Matching on Different Variables

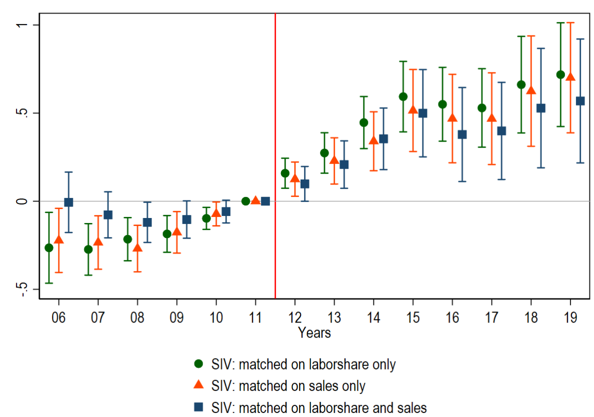
In the main text, I argue that matching on trends in labor share and sales is crucial to create synthetic firms that follow similar economic trajectories. As evidence, I report that pre-trends disappear also in the unmatched outcomes of employment and payroll. Figure D.9 shows that the trends do not disappear when I match only on labor share or only on sales.

Figure D.9a shows event-study estimates on firms' employment when I match on labor share only, sales only, and when I match on both jointly. Significant pre-trends remain in the pre-period when matching on single variables, which could raise concerns for the empirical design. Notably, the estimated effects in the post-period are similar across all matching types. This occurs because the estimated effects in the post-period are substantially larger than the magnitude of residual differential trends after the matching step. Figure D.9b repeats the analysis for payroll and finds the same results.

Figure D.9: Effects of matching on different variables on unmatched outcomes



(a) Number of employees



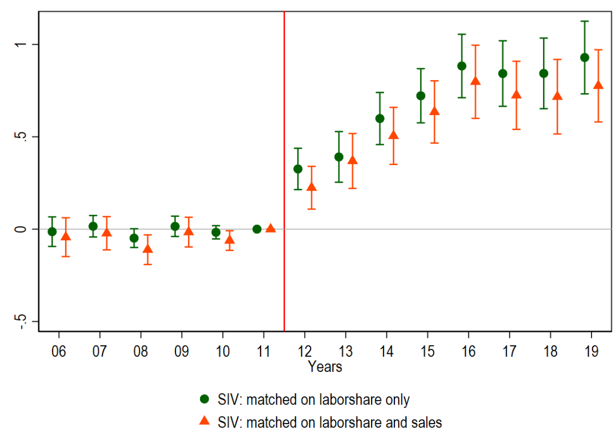
(b) Payroll

Notes: This figure presents event-study estimates using the SIV design, where synthetic control weights are calculated using different matching metrics. Green circles represent estimates from matching on trends in log labor share only, orange triangles show results from matching on trends in log sales only, and blue rectangles display the preferred estimates, which match on trends in both labor share and sales simultaneously.

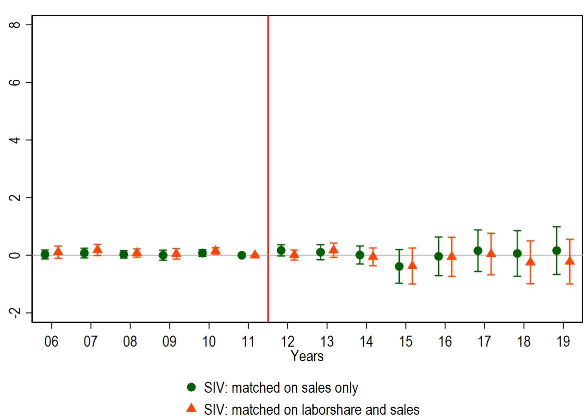
D.3 Robustness to Improving Pre-Treatment Fit

Figure D.10a shows the estimated effects when I match to improve pre-treatment fit for each specific outcome. Panel A shows upstream exposure effects on labor share when I match only on labor share (which improves pre-treatment fit) versus the default of matching on labor share and sales. Panels B and C show the estimated effects of D1 and D2 exposures, respectively, on sales when I match on sales only. Improving the pre-treatment fit does not change the results, providing evidence against concerns about under-fitting in the joint matching approach.

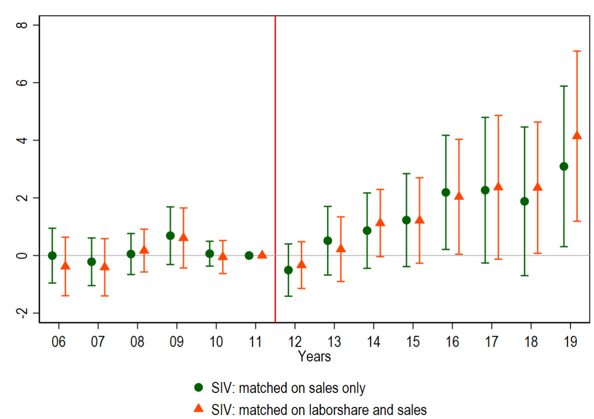
Figure D.10: Improving training fit: labor share and sales



(a) Labor share



(b) Sales: D1 exposure



(c) Sales: D2 exposure

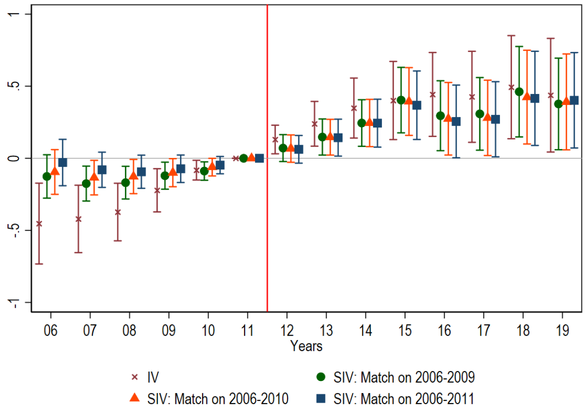
Notes: This figure shows event-study estimates of the effects of upstream exposure on labor share and downstream exposure on sales, comparing the preferred SIV estimator (matching on both labor share and sales) with estimators that match only on trends in the outcome: labor share in Panel A and sales in Panel B. Matching solely on the outcome of interest improves the training fit, but because the baseline fit is already strong, this refinement produces no detectable change in the post-period estimates.

D.4 Robustness to Different Training Periods

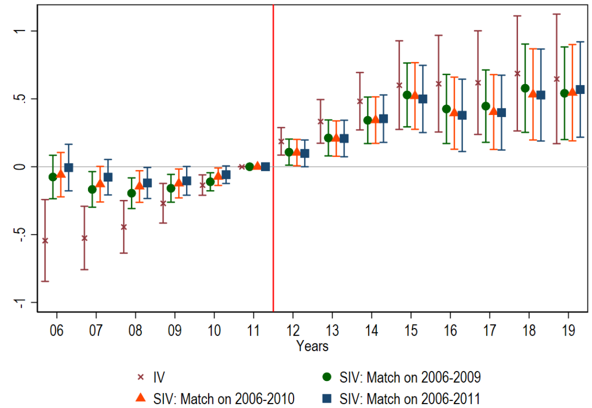
To test for overfitting, I perform back-testing by matching on trends in 2006–2009 and 2006–2010 instead of the default 2006–2011 period. This reduces the amount of information I match on but provides a visual check of sensitivity to particular matching periods.

Figure D.11 shows robustness checks of the main results using the travel-based instrument and matching on labor share and sales. Panels A, B, and C show effects of upstream exposure on employment, payroll, and labor share, respectively, while Panels D and E show effects of D1 and D2 exposures on sales. In each panel, I plot four sets of event-study estimates: baseline IV as a comparison, and SIV estimates using different training periods. SIV finds consistent signals across all variables even using the shorter 2006–2009 period. Hence, results are robust to matching on different pre-periods.

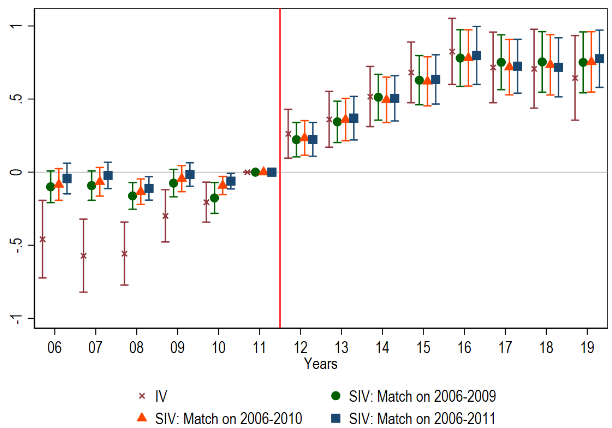
Figure D.11: Robustness to different training periods



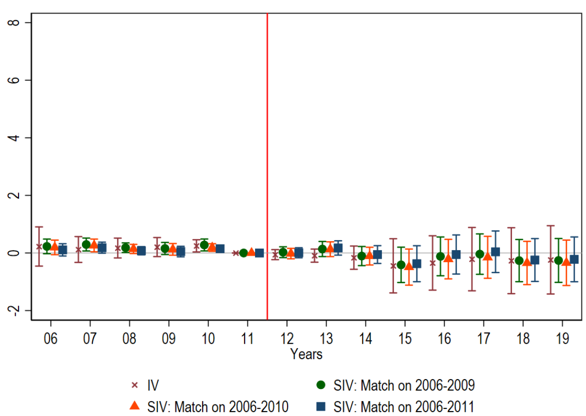
(a) Number of employees



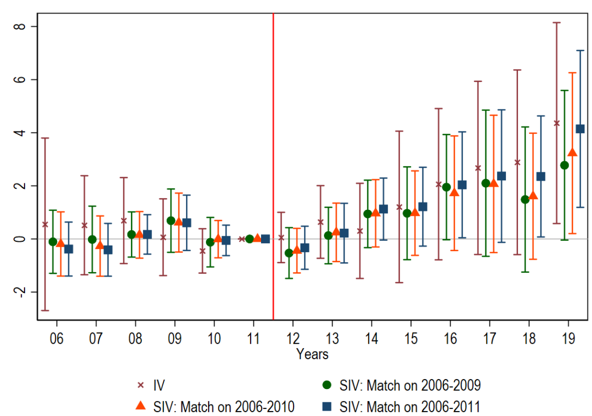
(b) Payroll



(c) Labor share



(d) Sales: D1-exposure



(e) Sales: D2-exposure

Notes: This figure shows event-study estimates from the SIV design where synthetic control weights are calculated using different training periods. The preferred specification uses the full pre-period (2006–2011) for training. Alternative specifications use truncated training periods of 2006–2009 and 2006–2010. All specifications outperform the reduced-form IV estimates, which are included for comparison.

D.5 2SLS estimates

The main text focuses on structural equations linking upstream exposure to labor share and downstream exposures to sales. Here I extend the analysis to employment and payroll, and test robustness across multiple specifications. I estimate effects of upstream and downstream exposures on all four outcomes, compare results for all manufacturing firms versus large firms only, and employ alternative instruments. Six tables report these checks, one for each instrument.

Table Structure

Each table follows an identical structure. Panel A includes all manufacturing firms; Panel B restricts to firms with 50+ baseline employees. Columns 1–3 report employment effects, 4–6 payroll, 7–9 labor share, and 10–12 sales. Within each outcome group, the first column isolates upstream exposure, the second isolates downstream exposures, and the third includes all three forces simultaneously.

I present these specifications for transparency about how sample restrictions affect results. The comparison between small and large firms helps identify where informality creates first-order bias, though I do not claim causal effects for small firms given their high informality rates.

Distance Instrument Results

Table D.3 presents results using the distance instrument. Other instruments yield similar patterns.

Employment (Columns 1–3): Upstream exposure increases employment for both large firms and the full sample, with effects significant at 5%. The larger sample of all firms yields more precise estimate, but point estimates remain comparable across samples.

Payroll and Labor Share (Columns 4–9): Results mirror employment effects. Upstream exposure increases labor demand across both samples, with economically and statistically similar magnitudes whether or not I control for downstream exposures.

Sales (Columns 10–12): Here the samples diverge sharply. For large firms, D1 exposure has no effect while D2 exposure significantly increases sales, a pattern robust to controlling for upstream exposure. For all firms, D1 effect remains null but D2 effect becomes negative (though insignificant). This negative D2 effect would violate the model's logic: it implies intermediates substitute more easily with labor than with each other, contradicting the positive upstream exposure and null D1 exposure effects. This pattern suggests informality biases the sales estimates for small firms.

D.5.1 Why Informality Matters: A Mechanism

The substantial role of informality in attenuating estimated effects on sales warrants examination. I propose that the primary mechanism operates through coordinated adjustments across multiple margins of informality. Building on Gulek (2024), which documents that firms substituted informal for formal labor following Syrian refugee arrivals (who lacked work permits and could only work informally), I explore implications for firms' transaction patterns.

Firms face fundamental accounting constraints when managing formal and informal transactions. Formal revenues generate documented cash flows through the banking system and must finance formal costs requiring official documentation. Similarly, informal revenues (undocumented cash transactions) provide the liquidity necessary to finance informal costs, particularly off-the-books wage payments.

When firms reduce formal labor costs and increase informal labor costs in response to cheaper informal labor, maintaining consistent accounting records becomes challenging. A decrease in formal costs without corresponding adjustments could create implausible profit margins triggering tax authority scrutiny. Firms therefore face pressure to coordinate adjustments across multiple margins.

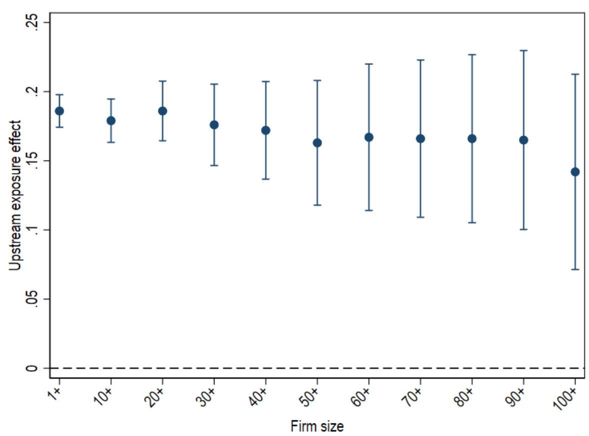
This coordination manifests through a cascade of informality. As firms reduce formal labor costs, they shift sales from formal to informal channels to generate undocumented cash for informal wage payments. However, this creates new inconsistencies: lower formal sales and labor costs while formal intermediate input purchases remain unchanged would appear suspicious. Firms therefore also shift some intermediate input purchases from formal to informal suppliers, completing coordination across labor, sales, and intermediate input margins.

This mechanism implies that informality in labor markets begets informality in product markets, which begets informality in input purchases. The coordination is driven by the need to maintain plausible accounting relationships rather than technological considerations. Firms may be relatively indifferent between formal and informal suppliers for certain inputs, making this adjustment feasible without significant productivity losses.

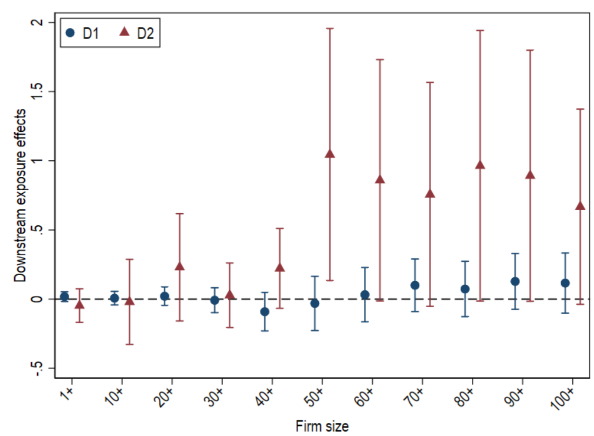
The Syrian refugee shock, by expanding informal labor supply, may have triggered these coordinated adjustments throughout small firms' operations. This explains why measured impacts on formal sales are biased for small firms: a portion of the sales response occurs through informal channels that escape measurement in official data.

Supporting evidence for the claim that D2-exposure effects on sales are biased for small firms due to informality appears in Figure D.12, which shows how upstream and downstream exposure effects change across firms of different sizes. There is a clear jump in estimated treatment effects at the 50-employee threshold, precisely when firms become significantly more formal due to increased government scrutiny.

Figure D.12: Estimated effects of trade exposures across different firm sizes



(a) Upstream Exposure Effects on Labor share



(b) Downstream Exposure Effects on Sales

Notes: the 95% confidence intervals of SIV estimates, where the training period is 2006–2010, weights are calculated by matching on the trends in labor share and sales, and travel distance is used as an instrument. Each column shows the estimated effects using a different sample of mmanufacturing firms, where "x+" denotes firms with at least x employees in 2010.

Table D.3: Effects of Trade Exposures on Firms in Non-Host Regions (IV: Distance)

	Number of employees			Payroll			Labor Share			Sales		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: All sizes												
U	0.056***		0.050***	0.077***		0.073***	0.186***		0.204***	-0.013*		-0.014
	(0.006)		(0.007)	(0.006)		(0.007)	(0.006)		(0.007)	(0.007)		(0.009)
D1		-0.101***	-0.026		-0.129***	-0.019		-0.221***	0.084***		0.018	-0.003
		(0.015)	(0.018)		(0.016)	(0.020)		(0.015)	(0.017)		(0.018)	(0.023)
D2		0.048	0.038		0.033	0.019		0.108**	0.069*		-0.047	-0.044
		(0.060)	(0.059)		(0.057)	(0.056)		(0.042)	(0.041)		(0.062)	(0.063)
Panel B: 50+ employees												
U	0.084**		0.077	0.114***		0.086	0.163***		0.193***	0.012		-0.017
	(0.037)		(0.050)	(0.039)		(0.054)	(0.023)		(0.031)	(0.038)		(0.053)
D1		-0.128	-0.013		-0.205*	-0.077		-0.217***	0.073		-0.031	-0.056
		(0.107)	(0.147)		(0.120)	(0.167)		(0.058)	(0.070)		(0.100)	(0.140)
D2		0.378	0.354		0.340	0.313		-0.577**	-0.637**		1.045**	1.051**
		(0.361)	(0.361)		(0.380)	(0.380)		(0.292)	(0.283)		(0.465)	(0.466)

Notes: Sample is restricted to manufacturing firms that report positive sales throughout 2006–2019. Panel A shows the SIV estimates for firms from all sizes. Panel B shows the results for firms with 50+ employees in 2010.There are 19155 firms in Panel A and 1224 firms in Panel B. U denotes the upstream exposure on firms who directly or indirectly buy from the host regions. D1 is the downstream exposure effect capturing cross-price elasticity between different intermediate goods. D2 is the downstream exposure on firms who directly or indirectly sell to the host region firms. Standard errors are clustered at the firm level.

Table D.4: Effects of Trade Exposures on Firms in Non-Host Regions (IV: Language)

	Number of employees			Payroll			Labor Share			Sales		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: All sizes												
U	0.054***		0.042***	0.076***		0.067***	0.173***		0.187***	-0.002		-0.009
	(0.007)		(0.011)	(0.008)		(0.012)	(0.008)		(0.012)	(0.009)		(0.011)
D1		-0.117***	-0.060		-0.138***	-0.047		-0.191***	0.062		-0.026	-0.038
		(0.028)	(0.037)		(0.030)	(0.040)		(0.026)	(0.038)		(0.025)	(0.032)
D2		-0.189**	-0.166*		-0.306***	-0.270***		-0.182	-0.082		-0.282**	-0.287**
		(0.089)	(0.088)		(0.104)	(0.100)		(0.128)	(0.105)		(0.129)	(0.129)
Panel B: 50+ employees												
U	0.022		-0.142	0.047		-0.143	0.094***		0.017	0.024		-0.097
	(0.043)		(0.166)	(0.045)		(0.191)	(0.034)		(0.151)	(0.038)		(0.091)
D1		-0.346	-0.485		-0.453	-0.594		-0.330	-0.313		-0.222	-0.318
		(0.524)	(0.564)		(0.581)	(0.645)		(0.350)	(0.506)		(0.301)	(0.313)
D2		0.079	0.188		-0.188	-0.079		-0.706	-0.719		0.439	0.514
		(0.838)	(0.583)		(0.920)	(0.641)		(0.577)	(0.501)		(0.553)	(0.446)

Notes: Sample is restricted to manufacturing firms that report positive sales throughout 2006–2019. Panel A shows the SIV estimates for firms from all sizes. Panel B shows the results for firms with 50+ employees in 2010.There are 19155 firms in Panel A and 1224 firms in Panel B. U denotes the upstream exposure on firms who directly or indirectly buy from the host regions. D1 is the downstream exposure effect capturing cross-price elasticity between different intermediate goods. D2 is the downstream exposure on firms who directly or indirectly sell to the host region firms. Standard errors are clustered at the firm level.

Table D.5: Effects of Trade Exposures on Firms in Non-Host Regions (IV: Host region indicator)

	Number of employees			Payroll			Labor Share			Sales		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: All sizes												
U	0.049***		0.048***	0.068***		0.072***	0.158***		0.177***	-0.004		0.007
	(0.007)		(0.008)	(0.007)		(0.009)	(0.007)		(0.009)	(0.008)		(0.010)
D1		-0.066***	-0.007		-0.073***	0.017		-0.135***	0.084***		0.032	0.040
		(0.018)	(0.021)		(0.019)	(0.023)		(0.017)	(0.020)		(0.021)	(0.027)
D2		-0.063	-0.078		-0.092	-0.116		0.100	0.042		-0.297***	-0.299***
		(0.083)	(0.082)		(0.099)	(0.096)		(0.080)	(0.078)		(0.105)	(0.105)
Panel B: 50+ employees												
U	0.067*		0.029	0.086**		0.046	0.067**		0.104***	0.076*		0.027
	(0.038)		(0.053)	(0.040)		(0.054)	(0.028)		(0.037)	(0.040)		(0.057)
D1		-0.140	-0.100		-0.174*	-0.112		-0.069	0.074		-0.125	-0.088
		(0.099)	(0.134)		(0.102)	(0.137)		(0.055)	(0.072)		(0.094)	(0.132)
D2		0.163	0.142		0.119	0.087		-0.303	-0.375		0.622*	0.604*
		(0.336)	(0.343)		(0.350)	(0.361)		(0.245)	(0.243)		(0.345)	(0.353)

Notes: Sample is restricted to manufacturing firms that report positive sales throughout 2006–2019. Panel A shows the SIV estimates for firms from all sizes. Panel B shows the results for firms with 50+ employees in 2010.There are 19155 firms in Panel A and 1224 firms in Panel B. U denotes the upstream exposure on firms who directly or indirectly buy from the host regions. D1 is the downstream exposure effect capturing cross-price elasticity between different intermediate goods. D2 is the downstream exposure on firms who directly or indirectly sell to the host region firms. Standard errors are clustered at the firm level.

Table D.6: Effects of Trade Exposures on Firms in Non-Host Regions (IV: Distance*Skill)

	Number of employees			Payroll			Labor Share			Sales		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: All sizes												
U	0.070***		0.060***	0.097***		0.088***	0.234***		0.251***	-0.015*		-0.016
	(0.007)		(0.009)	(0.008)		(0.010)	(0.008)		(0.010)	(0.009)		(0.012)
D1		-0.134***	-0.048**		-0.170***	-0.043*		-0.285***	0.077***		0.019	-0.004
		(0.019)	(0.023)		(0.020)	(0.025)		(0.019)	(0.022)		(0.022)	(0.029)
D2		-0.024	-0.066*		0.018	-0.044		0.165***	-0.014		-0.066	-0.055
		(0.034)	(0.038)		(0.036)	(0.040)		(0.046)	(0.047)		(0.051)	(0.052)
Panel B: 50+ employees												
U	0.097**		0.080	0.135***		0.093	0.185***		0.233***	0.026		-0.033
	(0.045)		(0.062)	(0.048)		(0.066)	(0.030)		(0.039)	(0.047)		(0.065)
D1		-0.170	-0.056		-0.266*	-0.133		-0.205***	0.131		-0.127	-0.174
		(0.127)	(0.173)		(0.143)	(0.197)		(0.070)	(0.087)		(0.116)	(0.162)
D2		-0.002	-0.091		-0.001	-0.104		-0.232	-0.491**		0.240	0.276
		(0.227)	(0.227)		(0.237)	(0.239)		(0.186)	(0.192)		(0.234)	(0.245)

Notes: Sample is restricted to manufacturing firms that report positive sales throughout 2006–2019. Panel A shows the SIV estimates for firms from all sizes. Panel B shows the results for firms with 50+ employees in 2010.There are 19155 firms in Panel A and 1224 firms in Panel B. U denotes the upstream exposure on firms who directly or indirectly buy from the host regions. D1 is the downstream exposure effect capturing cross-price elasticity between different intermediate goods. D2 is the downstream exposure on firms who directly or indirectly sell to the host region firms. Standard errors are clustered at the firm level.

Table D.7: Effects of Trade Exposures on Firms in Non-Host Regions (IV: Language*Skill)

	Number of employees			Payroll			Labor Share			Sales		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: All sizes												
U	0.070***		0.053***	0.098***		0.086***	0.213***		0.232***	0.005		-0.006
	(0.010)		(0.014)	(0.011)		(0.016)	(0.011)		(0.016)	(0.011)		(0.014)
D1		-0.143***	-0.074		-0.165***	-0.054		-0.218***	0.081		-0.039	-0.047
		(0.035)	(0.047)		(0.037)	(0.052)		(0.031)	(0.050)		(0.032)	(0.041)
D2		-0.101	-0.109		-0.174*	-0.188*		-0.159	-0.197*		-0.100	-0.099
		(0.088)	(0.088)		(0.101)	(0.100)		(0.102)	(0.104)		(0.130)	(0.130)
Panel B: 50+ employees												
U	0.015		-0.150	0.044		-0.131	0.100**		0.051	0.033		-0.105
	(0.054)		(0.115)	(0.056)		(0.129)	(0.044)		(0.110)	(0.047)		(0.069)
D1		-0.360	-0.485		-0.438	-0.547		-0.257	-0.214		-0.285	-0.373*
		(0.417)	(0.380)		(0.433)	(0.424)		(0.233)	(0.355)		(0.243)	(0.218)
D2		0.270	0.444		-0.058	0.094		-0.619	-0.678*		0.599	0.721
		(0.637)	(0.501)		(0.651)	(0.515)		(0.398)	(0.371)		(0.487)	(0.458)

Notes: Sample is restricted to manufacturing firms that report positive sales throughout 2006–2019. Panel A shows the SIV estimates for firms from all sizes. Panel B shows the results for firms with 50+ employees in 2010.There are 19155 firms in Panel A and 1224 firms in Panel B. U denotes the upstream exposure on firms who directly or indirectly buy from the host regions. D1 is the downstream exposure effect capturing cross-price elasticity between different intermediate goods. D2 is the downstream exposure on firms who directly or indirectly sell to the host region firms. Standard errors are clustered at the firm level.

Table D.8: Effects of Trade Exposures on Firms in Non-Host Regions (IV: Host-region indicator * Skill)

	Number of employees			Payroll			Labor Share			Sales		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: All sizes												
U	0.060***		0.058***	0.083***		0.088***	0.192***		0.216***	-0.003		0.010
	(0.008)		(0.010)	(0.009)		(0.011)	(0.009)		(0.011)	(0.010)		(0.013)
D1		-0.085***	-0.012		-0.093***	0.017		-0.167***	0.104***		0.040	0.053
		(0.022)	(0.027)		(0.023)	(0.029)		(0.021)	(0.026)		(0.027)	(0.034)
D2		-0.016	-0.054		-0.025	-0.083		0.093	-0.050		-0.200**	-0.207**
		(0.080)	(0.080)		(0.094)	(0.093)		(0.078)	(0.078)		(0.098)	(0.100)
Panel B: 50+ employees												
U	0.078*		0.035	0.098**		0.056	0.074**		0.128***	0.092**		0.019
	(0.045)		(0.063)	(0.047)		(0.066)	(0.033)		(0.045)	(0.046)		(0.069)
D1		-0.175	-0.127		-0.209*	-0.132		-0.056	0.122		-0.192*	-0.165
		(0.114)	(0.157)		(0.116)	(0.163)		(0.061)	(0.082)		(0.103)	(0.151)
D2		0.095	0.061		0.045	-0.009		-0.321	-0.446		0.606*	0.587
		(0.371)	(0.393)		(0.386)	(0.414)		(0.272)	(0.279)		(0.364)	(0.387)

Notes: Sample is restricted to manufacturing firms that report positive sales throughout 2006–2019. Panel A shows the SIV estimates for firms from all sizes. Panel B shows the results for firms with 50+ employees in 2010.There are 19155 firms in Panel A and 1224 firms in Panel B. U denotes the upstream exposure on firms who directly or indirectly buy from the host regions. D1 is the downstream exposure effect capturing cross-price elasticity between different intermediate goods. D2 is the downstream exposure on firms who directly or indirectly sell to the host region firms. Standard errors are clustered at the firm level.

E Additional Counterfactual Estimates

Comparison between Adana and Antalya

In the main text, I argue that a host region's centrality is the most informative factor in determining the magnitude of trade spillovers from immigration. To strengthen this argument, I compare two cities, Adana and Antalya, that share similar population sizes and Domar weights but differ significantly in their economic connectedness due to their industrial compositions.

Table E.9 presents baseline statistics for these cities. In 2010, Adana had a population of 2.11 million (5th largest in Turkey), while Antalya had 2.04 million (6th largest). Their Domar weights were similar: 2.48% for Adana (7th highest) and 2.70% for Antalya (6th highest).

Despite these similarities, the cities exhibit marked differences in industrial structure: Adana serves as an agricultural hub, whereas Antalya's economy centers on tourism and services. These distinctions manifest in their cost-based and sales-based Bonacich centrality measures, with Adana's measures being 1.7 and 1.4 times larger than Antalya's, respectively. This difference in economic centrality translates directly into spillover effects: the average spillover wage effect from Adana is 2.4 times greater than that from Antalya.

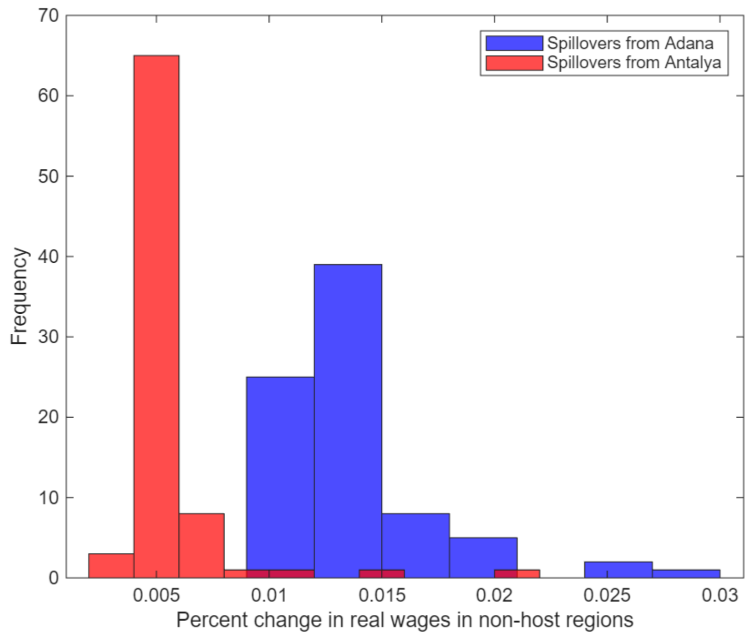
Table E.9: Summary Statistics for Adana and Antalya

	Adana	Antalya
Population (in millions)	2.11	2.04
Domar weight	0.025	0.027
Cost-based centrality: $\tilde{\Psi}1$	10.94	6.54
Sales-based centrality: $\Psi 1$	2.21	1.60
Spillover effect on real wages	1.37%	0.56%

Source: Authors' calculations

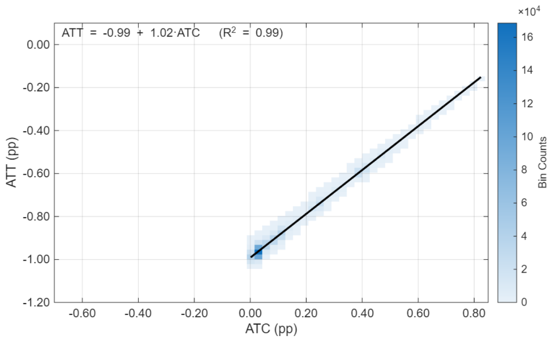
A potential concern with the mean spillover difference presented in Table E.9 is that it could be driven by a small number of outliers. To address this issue, Figure E.13 displays the distribution of spillover effects resulting from a 1% immigration shock to each city. The histograms reveal that the spillover distributions for Adana and Antalya barely overlap, with the minimum spillover effect from Adana exceeding the 95th percentile of spillover effects from Antalya. This stark separation in distributions confirms that the difference in spillover effects is systematic rather than driven by outliers.

Figure E.13: Histogram of real wage changes in the non-host regions

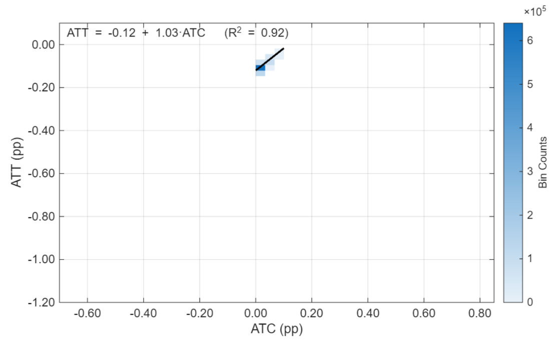


Notes: This figure shows the spillover effects from two counterfactuals: a 1% increase in labor supply in Adana and Antalya. Adana and Antalya share similar population sizes and Domar weights but differ significantly in their economic connectedness due to their industrial compositions. Adana is more central as it is an Agricultural hub, while Antalya has a more tourism and services based economy.

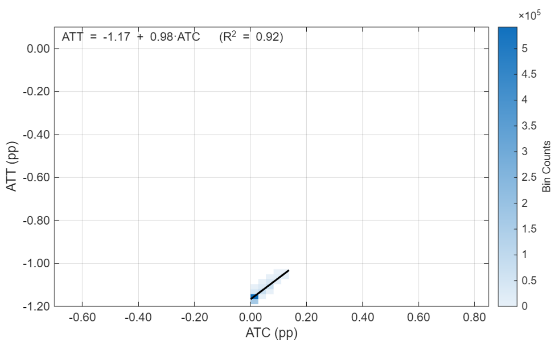
Figure E.14: ATT and ATC estimates of real and nominal wages across simulations with different elasticity parameters



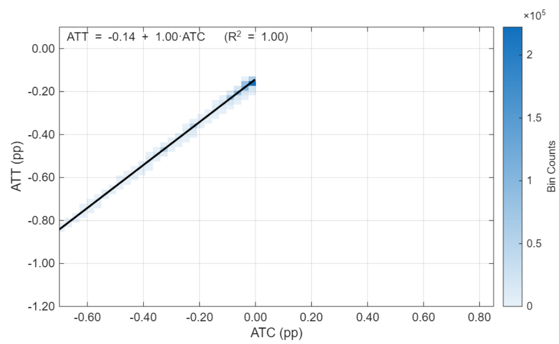
(a) Real wages: $\sigma_u = 0.79$, $\sigma_L = 1.06$



(b) Real wages: $\sigma_u = 10$, $\sigma_L = 1$



(c) Nominal wages: $\sigma_u = 0.79$, $\sigma_L = 1.06$



(d) Nominal wages: $\sigma_u = 10$, $\sigma_L = 1$

Notes: Each panel shows results from one million simulations under different elasticity parameters. In each simulation, 10 randomly selected regions receive a 1% labor supply increase. Changes in nominal wages and prices are calculated by solving the system of linear equations in Theorem 1. Real wage changes are defined as $d\ln w_{real} = d\ln w - b * d\ln p$, where $d\ln w$ is the vector of nominal wage changes, b is the $R \times N$ matrix of final expenditure shares, and $d\ln p$ is the $N \times 1$ vector of price changes. ATT is the average wage change in treated regions; ATC is the average change in control regions. Each panel displays binscatter plots, the linear fit, and the bivariate regression equation.