Formal Effects of Informal Labor and Work Permits Evidence from Venezuelan Refugees in Colombia

Dany Bahar Isabel Di Tella Ahmet Gulek*

August 11, 2025

Most recent draft here.

Abstract

We analyze the Venezuelan refugee crisis in Colombia to separately identify effects of informal immigration and work permit policies on labor markets. Using Synthetic Instrumental Variables and triple difference-in-differences designs, we find that the informal labor supply shock displaced native workers in both informal and formal sectors, indicating high substitutability between worker types (elasticity ≈ 11). Work permits reduced competition in the informal sector while increasing it in the formal sector, creating 24,440 new formal jobs and approximately $43 million in annual tax revenue. Results suggest work permits create productivity spillovers through reduced skill mismatch, providing economic rationale for immigrant integration policies.

JEL Classification: D22, J21, J46.

Keywords: Immigration, Refugee crises, Work permits, Informality

*Bahar: Brown University and Harvard Growth Lab (email: dany_bahar@brown.edu); Di Tella: PhD student, MIT Department of Economics (e-mail: iditella@mit.edu); Gulek: Postdoctoral scholar, Immigration and Policy Lab at Stanford (e-mail: agulek@stanford.edu). We are thankful to Daron Acemoglu, Josh Angrist, David Autor, Abhijit Banerjee, Martin Beraja, Amy Finkelstein, Simon Jaeger, participants in The Sixth World Labor Conference hosted by SOLE, the MIT Labor Lunch and Trade Lunch. We acknowledge support from the George and Obie Shultz Fund at MIT.

1 Introduction

Forced immigration episodes have long been used as quasi-experimental settings to study the effects of immigration on labor markets. Examples of such episodes include the Mariel Boatlift (Card, 1990), the Algerian war of independence (Hunt, 1992), the Yugoslav wars (Angrist and Kugler, 2003), and the Syrian refugee crisis (Del Carpio and Wagner, 2015). However, these episodes often entail two distinct shocks occurring at once: the arrival of immigrants and subsequent government policies that either facilitate or restrict immigrants' integration. They include, but are not limited to, freedom of movement, right to work, access to public services, and pathways to permanent residency. Naturally, the overall effect of these episodes on the host country's labor markets likely depends on which rights are granted. However, existing research typically conflates these distinct policy interventions with the initial immigration shock. In this paper, we present a theoretical analysis formalizing the impact of immigrants with and without work permits and empirical evidence documenting these forces in the context of Venezuelan migration to Colombia.

Colombia is an ideal setting to study the separate effects of immigration and work permit policies because it is the only country to receive a large number of irregular migrants and only later grant them work permits en masse.[1] By 2023, Colombia hosted 2.5 million forcibly displaced Venezuelans, one of the largest refugee-like populations in the world. Venezuelans began arriving in 2015, mostly without permission to work. Over the course of several waves beginning in 2017, Colombia provided Venezuelans with a migratory status that included work permits. The gradual provision of these work permits provides quasi-experimental variation that allows us to identify the labor market effects of both (1) a largely informal labor supply shock and (2) the subsequent provision of work permits.

We start by proposing an equilibrium model that incorporates two key aspects of labor markets in developing economies: firms can utilize both informal and formal labor in production, and natives can work either for firms in salaried positions or become self-employed (approximately half of all workers have non-salaried work according to our data). In this model a positive informal labor supply shock, such as the arrival of immigrants without a legal status, necessarily reduces natives' wages and salaried employment in the informal sector. However, higher overall (natives and immigrants) informal salaried employment has two competing effects in the formal sector: it makes formal workers more productive through Q-complementarity, while also creating competition against formal employees subject to diminishing returns to labor. When informal and formal workers are gross substitutes in pro-

[1] In contrast, Syrian refugees in Turkey do not have permits to this day (Gulek, 2023), while Ukrainian refugees received permits almost immediately in all EU countries (European Commission, n.d.).

duction, the overall effect is to incentivize firms to shift toward greater informality. When informal wages fall, natives who can transition to non-salaried positions without substantial pay cuts choose to do so. This transition both limits the decline in informal wages and amplifies native displacement in salaried employment. A testable implication is that total informal employment may remain relatively stable while natives are systematically replaced in salaried positions.

In this framework, work permits shift some labor supply from the informal to the formal sector, reducing competition in the informal sector while increasing direct competition in the formal sector. Consequently, the first-order effects are that natives' wages and salaried employment increase in the informal sector but decrease in the formal sector. However, the model highlights a crucial equilibrium effect on total formal employment in the economy. When informal and formal workers are highly substitutable, the increase in informal wages causes firms to substitute away from informal workers toward formal workers, effectively increasing total formal employment despite the direct displacement of formal native workers.

Next, we empirically test the model's implications. Our key empirical challenges are threefold. First, immigrants choose where to locate based on economic conditions, leading to the standard endogeneity problem. Second, even regions that receive immigrants for plausibly exogenous reasons, such as proximity to the border, may follow different economic trajectories to regions that don't receive them. Third, regions with more immigrants tend to receive more work permits, creating spatial correlation between the two. To address these challenges, we employ two distinct research designs that separately identify the labor market effects of (1) a mostly informal labor supply shock and (2) granting work permits to irregular migrants.[2]

To study the effects of the mostly informal labor supply shock, we employ a shift-share design where travel distance between Colombian and Venezuelan cities operates as the share and the total number of refugees in Colombia in a given year acts as the shift. We further relax the share-exogeneity assumption embedded in our design (Goldsmith-Pinkham et al., 2020) by using the Synthetic Instrumental Variables method (SIV) developed in Gulek and Vives-i Bastida (2023). This method uses synthetic controls to account for the empirical fact that border regions follow different economic trajectories to other regions.

We find that informal immigration displaced native salaried workers in both the informal and formal sectors. A 1 percentage point (pp) increase in the refugee/native ratio decreases native informal salaried employment by 0.26 pp and formal salaried employment by 0.16

[2]We characterize this as a "mostly informal" labor supply shock because it includes a smaller component of return migration by Colombian citizens from Venezuela, who retain legal access to formal employment. However, Venezuelan migrants constitute the overwhelming majority of this population movement.

pp.[3] The former result is predicted by the model, while the latter suggests that informal and formal labor are highly substitutable in production, with an estimated elasticity of substitution of approximately 11. Moreover, we find that the immigration shock *increases* non-salaried employment, especially for men without high-school degrees, as predicted by the model. This transition is numerically large enough that total informal employment rates remain unchanged in the aggregate.

To study the effects of work permits on natives, we exploit the fact that work permits only enable skilled immigrants to transition to the formal sector, while having no effect on the less-skilled. This additional margin of comparison allows us to employ a triple difference-in-differences design that partials out the region-time specific immigration shocks identified earlier.

We find that work permits cause natives to gain informal jobs and to lose formal jobs. This is consistent with the model, which rationalizes these effects as arising from reduced competition in the informal sector and increased competition in the formal sector. We further estimate the model using our empirical estimates and moments from the data, and use the model to quantify work permits' effect on formal job creation. We estimate that all work permits issued during our study period enabled 68,000 Venezuelans to transition from informal to formal jobs, creating a net increase of 24,440 formal jobs in the economy. Under the assumption that all these workers received the mean wage in the formal sector (despite being higher skilled than the average formal worker), we estimate that granting work permits generated approximately 43 million USD in annual tax revenue.

Finally, we investigate whether work permits facilitate better matches between migrants and employers, thereby generating productivity gains in the economy. We present suggestive evidence that work permits lead to wage increases for more exposed natives in formal, skill-intensive industries (such as finance) relative to less formal, less skill-intensive industries (such as construction and transportation). Our quadruple difference-in-differences design shows that this effect is unlikely to be driven by industry-, skill-sex-, region-, or time-specific trends, or any triple combination of such trends. These results are consistent with relatively higher-skilled refugees transitioning from low-skill to high-skill industries where they provide productivity gains beyond their labor supply. This finding suggests that granting refugees work permits may additionally generate productivity gains in the economy.

These findings contribute to several strands of literature. First, this paper comprehensively reexamines labor market responses to immigration shocks (e.g., Card, 2009; Glitz, 2012; Tumen, 2016; Borjas and Monras, 2017; Llull, 2018; Peri and Yasenov, 2019; Lebow,

[3] These estimates represent, based on average values at baseline, about a 3% drop in informal salaried employment (baseline of 9%) and less than a 1% drop in formal salaried employment (baseline of 18%).

2024). The primary contribution of our paper is, to the best of our knowledge, the first separate empirical estimates of the effects of the inflow of immigrants and subsequent policies. Researchers who conflate immigration shocks with their subsequent policies may systematically underestimate or overestimate policy effects due to confounding with the initial shock. For example, in our context we demonstrate that while the informal immigration shock displaces natives in both the informal and formal sectors, granting work permits to immigrants improves natives' prospects in the informal sector and increases the total number of formal jobs in the economy.

Second, we complement the literature studying the dynamics between informal and formal labor markets. Initial contributions in this field were largely theoretical (Rauch, 1991; Amaral and Quintin, 2006), while more recent efforts have focused on calibrating and estimating structural models (Bosch and Esteban-Pretel, 2012; Meghir et al., 2015; Ulyssea, 2018). One exception is Gulek (2023), who studies the labor market consequences of the Syrian refugee shock in Turkey. Similarly, he finds that the arrival of informal refugees causes native disemployment in both informal and formal salaried jobs, with low-skill native men transitioning to non-salaried employment. However, while he predicts the effects of work permits using counterfactuals, our setting allows us to empirically test and document these effects directly. We also provide the first estimate of the elasticity of substitution between informal and formal salaried workers in Colombia, which we estimate to be approximately 11.[4] This elasticity is the key parameter determining how immigrant arrivals with or without work permits affect labor markets and, more generally, the dynamics between informal and formal sectors. The large elasticity we document implies that policies making formal labor relatively cheaper or informal labor relatively more expensive can substantially incentivize firms toward greater formality.

Third, our results on the effects of work permits build on a literature studying the impact of different formalization policies in developing countries (Monteiro and Assunção, 2012; De Andrade et al., 2016; Rocha et al., 2018) and developed countries (Elias et al., 2025).[5] Two papers focusing on work permits in refugee crises are most relevant to our work. In policy work, Clemens et al. (2018) provide economic arguments explaining why refugee work permits can benefit both refugees and natives. Most similar to our setting, Bahar et al. (2021) study the effects of granting Venezuelan refugees work permits and find negative but

[4]The only other estimates for this elasticity around the world come from Gulek (2023) and Schramm (2014), who estimate this parameter to be approximately 10 in Turkey and 1.8 in Mexico, respectively.

[5]Elias et al. (2025) study the impact of regularizing 600,000 non-EU immigrants in Spain and find that while immigrant formal employment increased, native employment was unaffected, whereas we find a decrease for natives. Theoretically, this discrepancy can be rationalized if regularization in Spain caused a larger decrease in mismatch than in our setting. The productivity gains from reduced mismatch can offset the increased competition from formal sector labor supply.

negligible effects on the formal employment rate of Colombian workers. Our results differ for two reasons. Economically, while Bahar et al. (2021) examine employment as a whole, we demonstrate that salaried and non-salaried positions respond to different economic forces and show that work permit effects are significant specifically among salaried positions. Econometrically, while Bahar et al. (2021) employ a standard difference-in-differences approach that may conflate the effects of ongoing immigration and work permits, we isolate the effect of work permits using triple and quadruple difference-in-differences designs that partial out the regional impact of immigration.

Finally, our work fits in with several other studies that specifically examine the impact of Venezuelan immigration on Colombia's labor market, which is one of the largest displacement crises seen this century (Santamaria, 2020; Caruso et al., 2021; Penaloza-Pacheco, 2022; Lebow, 2022; Delgado-Prieto, 2024; García-Suaza et al., 2024). Using different identification strategies (primarily spatial IV-DiD), this literature has found inconclusive results on employment.[6] We demonstrate that these inconsistencies are driven by two factors: (i) misinterpreting differential trends as causal effects and (ii) overlooking men's transitions to non-salaried positions. We show that host regions follow different trajectories than other regions, implying that the share-exogeneity assumption embedded in these designs (Goldsmith-Pinkham et al., 2020) likely fails.[7] Our paper contributes to this literature by implementing the Synthetic IV estimator that adjusts for differential trends and endogenous treatment selection jointly, relaxing the share-exogeneity assumption (Gulek and Vives-i Bastida, 2023) and therefore providing more reliable estimates of causal effects. This approach reveals that natives lost salaried jobs in both informal and formal sectors across different skill-gender groups, while low-skill men transitioned into non-salaried positions. Our theoretical model rationalizes these effects, isolates the role of work permits, and enables us to provide the first estimate of the elasticity of substitution between informal and formal salaried workers in this context.

The rest of this paper is structured as follows. Section 2 provides background on Colombian labor markets and the Venezuelan refugee crisis. Section 3 introduces the economic framework and derives its theoretical predictions. Section 4 presents the effects of the im-

[6]Santamaria (2020); Caruso et al. (2021); Lebow (2022) find little to no impact on employment, while Delgado-Prieto (2024) finds decreases in formal employment but no effect on informal employment, which he attributes to the binding minimum wage in the formal sector. Our results show that natives lose salaried jobs in both informal and formal sectors across all skill groups.

[7]Similar concerns in this setting were initially raised by Lebow (2022) and Lebow (2024). However, the implemented fixes—such as applying synthetic controls without using an instrument (Penaloza-Pacheco, 2022), estimating trends and adding them as control variables (Delgado-Prieto, 2024), or controlling for regional-time trends in non-saturated regressions that estimate trends using post-treatment data (Lebow, 2022)— do not fully establish causal results, we claim.

migration shock. Section 5 analyzes the effects of work permits. Section 6 concludes.

2 Background and Data

2.1 The Venezuelan Refugee Crisis and the Colombian Policy Response

Venezuela began to experience out-migration in the early 2000's in response to President Chavez's political and economic reforms. Following his death in 2013, the economic and political crisis deepened. The first waves of refugees arrived in Colombia in late 2015, but their numbers remained small until mid-2016. As the economic crisis intensified in the following months, there was a substantial increase in Venezuelans seeking refuge in Colombia. Figure 1a shows how the number of Venezuelan refugees in Colombia has evolved over time. In 2022, Venezuelans made up approximately 5% of the Colombian population.

Figure 1b shows the distribution of the number of Venezuelans per 100 natives in Colombia at the department level, our unit of analysis, in 2019. As is standard in refugee crises, Venezuelans are more densely located in regions closer to the border (Caruso et al., 2021). The distance from the populous governorates of Venezuela strongly predicts the number of refugees per native in a given region, which constitutes one of the key building blocks of our identification strategy.

In response to the substantial Venezuelan migration inflows, the Colombian government established a temporary visa program called Permiso Especial de Permanencia (PEP), which granted Venezuelans authorization for formal employment and access to social security benefits. While we provide detailed documentation in Appendix B, we summarize the key features here for context. The PEP program was implemented through sequential waves. The initial two waves, launched in July 2017 and February 2018, targeted Venezuelans possessing proper documentation and a legally recognized migratory status. Consequently, these early waves exhibited strong positive selection, with eligibility predominantly restricted to highly skilled migrants. The first two waves of the PEP program resulted in the issuance of nearly 182,000 permits.

The third and largest wave of PEP was the byproduct of a nationwide registration of irregular migrants. As the number of undocumented, hence PEP-1 & 2 ineligible, Venezuelan immigrants was increasing, the Colombian government implemented a nationwide census known as RAMV between April and June 2018 to count and characterize the undocumented Venezuelan population in Colombia. Importantly, registration was not explicitly or implicitly related to the possibility of obtaining formal migratory status. RAMV identified 442,462

undocumented Venezuelan migrants in Colombia, belonging to 253,575 different households.

On July 25, 2018, just days before his term ended, President Juan Manuel Santos unexpectedly issued a decree allowing all undocumented Venezuelans listed in the RAMV to apply for a new round of PEP. Sixty-four percent, or about 280,000, of the total undocumented migrants registered in the RAMV received a PEP.

The differential documentation requirements between PEP waves 1 and 2 (formal documentation required) and PEP-3 (no documentation required) resulted in Venezuelans with varying skill levels accessing formal employment at different points in time. Figure 1c presents the number of formally employed Venezuelans across six education-gender categories. We distinguish three skill groups: less than high school education (low-skill), high school completion or above (middle-skill), and college completion or above (high-skill), with each education category further disaggregated by gender. Figure 1c demonstrates that high-skilled men began securing formal employment in the third quarter of 2017, coinciding with PEP waves 1 and 2, while middle-skilled men gained access to formal jobs only following PEP-3. Low-skilled men and women across all skill levels did not experience consistent or significant increases in formal job uptake.

In summary, the Venezuelan refugee crisis in Colombia began as an informal labor supply shock, but the subsequent granting of work permits enabled the highest-skilled Venezuelans to eventually transition into formal employment.

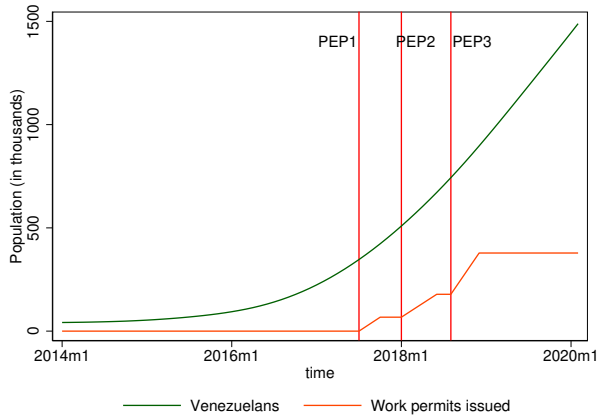
2.2 Labor market statistics

Information on informal and formal labor market outcomes for both Venezuelan and native workers is derived from the 2007–2019 Colombian Household Labor Force Surveys (*Encuesta Nacional de Hogares*—GEIH) conducted by the Colombian National Statistics Department (*Departamento Administrativo Nacional de Estadística*—DANE).[8] GEIH surveys approximately 20,000 households monthly and maintains representativeness at the department level across Colombia's 24 departments. Since 2013, GEIH has included questions regarding respondents' citizenship and place of residence one and five years prior to the survey date. Throughout this analysis, we use the term "natives" to refer to all non-Venezuelan residents unless otherwise specified.

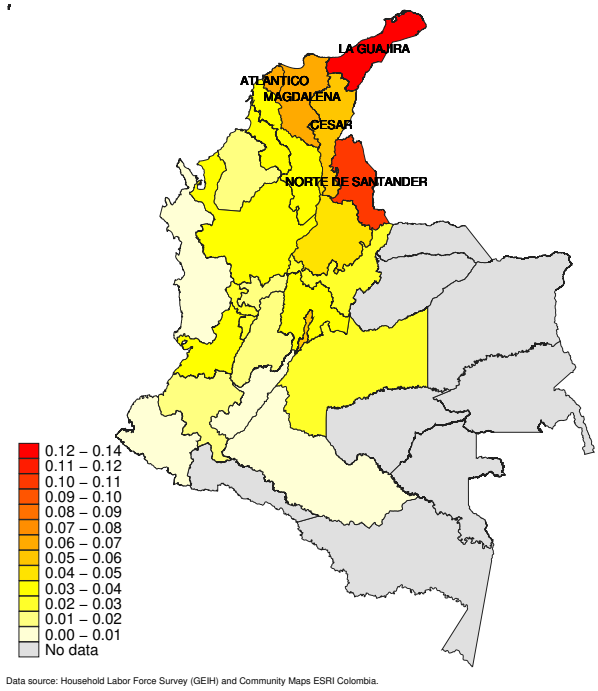
We construct three broad employment categories: salaried informal, salaried formal, and non-salaried workers. Salaried workers are individuals employed by private companies who receive regular wages. Non-salaried workers encompass all other employment arrangements, primarily self-employed individuals. We exclude government employment for analytical sim-

[8] This dataset is publicly available on DANE's website.

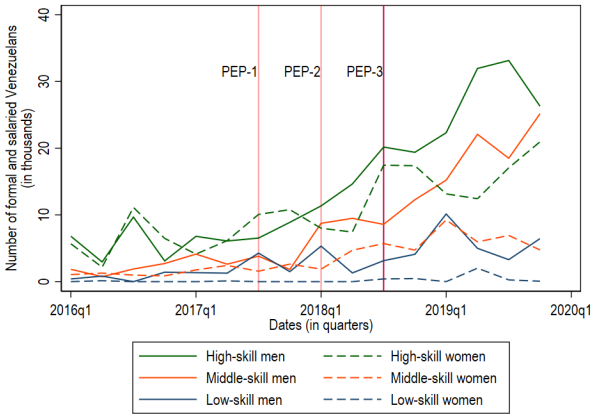
Figure 1: The Venezuelan Refugee Crisis in Colombia



(a) Immigration and Work permits



(b) Cross-sectional distribution of the Venezuelan to native ratio



(c) Formal salaried employment of Venezuelans

Notes: Authors' calculations.

plicity. This classification represents an economically meaningful distinction between jobs that depend primarily on firms' labor demand versus those driven by workers' labor supply decisions. This distinction is particularly relevant for interpreting migration impacts in developing country contexts where firms' labor demand alone does not capture the complete

labor market dynamics. For instance, a native worker who loses an informal salaried position and transitions to unpaid family work or self-employment remains classified as "employed" in the GEIH data. Consequently, examining aggregate native employment rates obscures important and economically meaningful compositional shifts within the labor force. To address this limitation, we analyze salaried and non-salaried employment categories separately. We do not differentiate between informal and formal status within non-salaried positions, since the overwhelming majority operate informally. Figure 2a illustrates employment rates across these three categories by gender and education groups. Salaried employment rates increase with educational attainment and are systematically higher for men. Non-salaried employment constitutes a substantial component of low-skilled men's overall employment.[9]

We distinguish between *formal* and *informal* employment based on respondent's self-reported social security coverage in GEIH. By law, Colombian employers must contribute to social security on their workers behalf.[10] This measure is a good predictor of formality for two reasons. First, there is no incentive for workers to misreport their insurance status. It is not illegal to work informally, only to employ informally. Second, the descriptive statistics on formal and informal employment using insurance status are consistent with the general knowledge on the informal sector (Ulyssea, 2020). Across regions and industries, the informality rate (defined as the ratio of employment that is informal) decreases with education. It is higher in less developed regions and in industries like agriculture, which are known to rely on informal labor.

Figures 2b and 2c show the formality rate of salaried workers across select industries and firm sizes. Although there is significant heterogeneity, formal and informal workers coexist everywhere. The formality rate across sectors ranges from 24% in private household work to 91% in financial intermediation. Moreover, formality rate increases as firms get bigger: ranging from 6% in firms with 1 employee, to 91% in firms with more than 101 employees. This relationship can be rationalized by the fact that larger firms are more visible and therefore more likely to be punished if hiring informally (Ulyssea, 2020).

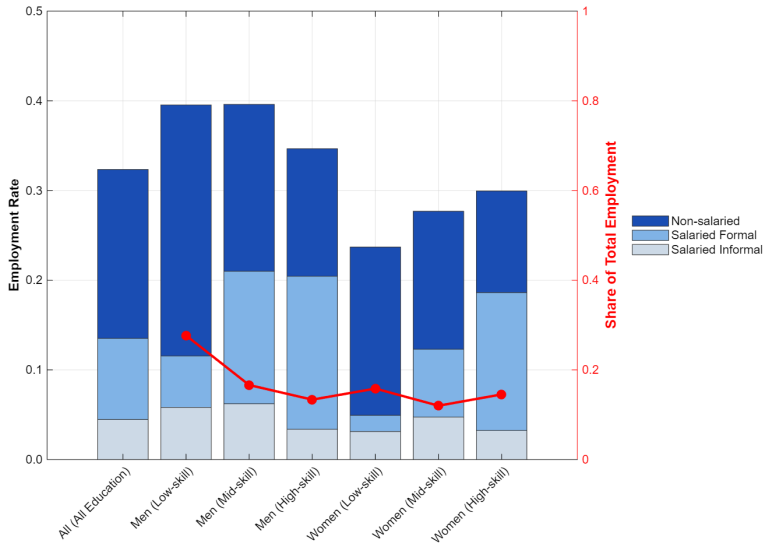
3 Model

We develop a theoretical framework that formalizes how informal labor supply increases and work permit policies affect native labor market outcomes in economies with substantial informality. Our framework incorporates a key institutional feature of developing country

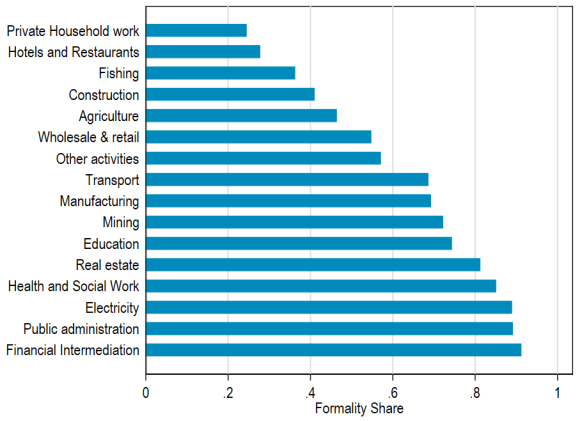
[9] Appendix Table A.1 provides more details regarding income and wages, and Table A.2 presents the distribution of employment types across our population and the six demographic subgroups.

[10] Law 100 of 1993

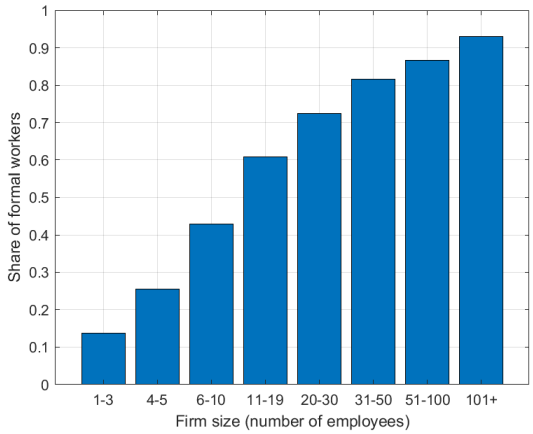
Figure 2: Summary Statistics of the Colombian Labor Market



(a) Employment rates of different sub-groups



(b) Informality across industries



(c) Informality across firm sizes

Notes: Authors' calculations using GEIH.

labor markets: workers can engage in either salaried employment (determined by firms' labor demand and workers' supply decisions) or non-salaried employment (determined solely by workers' optimization). A representative firm employs both informal and formal salaried workers in production. Given the widespread coexistence of informal and formal workers shown in Section 2.2, the representative firm assumption is a benign simplification that allows analytical tractability.

Labor Demand

The representative firm employs informal and formal salaried workers as the sole factors of production. Following Ulyssea (2018), we assume that while formal workers impose a constant marginal cost, hiring informal workers entails an increasing and convex expected cost. This cost structure reflects the greater enforcement risk faced by larger informal operations (De Paula and Scheinkman, 2011) and generates the empirically observed negative correlation between firm size and informal employment probability.

The firm maximizes profits by choosing optimal employment levels:

$$\max_{\ell_i,\ell_f} F(\ell_i,\ell_f) - \tau(\ell_i)w_i - w_f\ell_f, \tag{1}$$

where w_i and w_f denote informal and formal wages, respectively, and $\tau(\ell_i)$ captures the convex expected cost of hiring ℓ_i informal workers with $\tau'(\ell_i) > 0$ and $\tau''(\ell_i) > 0$. For analytical simplicity, we interpret w_f as the gross formal wage, abstracting from payroll taxes.

The production function takes the following constant elasticity of substitution (CES) form:

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

where θ represents Hicks-neutral productivity, $0 < \alpha < 1$ imposes decreasing returns to scale appropriate for short-run labor market adjustments, $\sigma = \frac{1}{1-\rho}$ denotes the elasticity of substitution between informal and formal labor, and η governs the relative productivity weights.[11]

Labor Supply

Workers are endowed with either formal or informal labor capabilities. Formal workers maximize utility subject to their budget constraint:

$$\max_{c,h_f} U(c) - \frac{\delta_f}{1+\delta_f}(h_f)^{\frac{1+\delta_f}{\delta_f}} \quad s.t. \quad c \leq h_f w_f, \tag{2}$$

where this quasi-linear specification yields a tractable formal labor supply function $\ell_f^* = w_f^{\delta_f}$ with δ_f representing the labor supply elasticity. The assumption that formal workers cannot access informal employment reflects institutional barriers that prevent formal sector workers

[11] This exact parametrization of the labor demand can also be found in Gulek (2023). Our theoretical contribution to his model is incorporating the labor supply choice between salaried and non-salaried jobs for informal workers.

from easily transitioning to informal work.[12]

Informal workers face a richer choice set, as they can work either in salaried positions for the representative firm or in non-salaried activities. Drawing on the home production literature (Gronau, 1977), we model non-salaried work as generating income $g(h_{ns})$ where $g(\cdot)$ is increasing and strictly concave, capturing diminishing returns to self-employment or informal entrepreneurial activities. The informal worker's optimization problem becomes:

$$\max_{c,h_i,h_{ns}} U(c) - \frac{\delta_i}{1+\delta_i}(h_i + h_{ns})^{\frac{1+\delta_i}{\delta_i}} \quad \text{s.t.} \quad c \leq h_i w_i + g(h_{ns}). \tag{3}$$

This specification implies that informal workers view salaried and non-salaried work as perfect substitutes in terms of disutility while facing different returns. The first-order condition for an interior solution yields $g'(h_{ns}) = w_i$, meaning workers allocate time to non-salaried activities until the marginal return equals the salaried informal wage. When immigration-induced competition reduces w_i, workers substitute toward non-salaried employment, with the magnitude of this transition inversely related to the curvature of $g(\cdot)$.

Equilibrium

Denoting $L^S_{i,s}(w_i)$, $L^S_{i,ns}(w_i)$, and $L^S_f(w_f)$ as informal salaried, informal non-salaried, and formal labor supply curves respectively, market clearing requires:

$$\begin{aligned} L^S_{i,s}(w_i) &= L^D_i(w_i, w_f), \\ L^S_f(w_f) &= L^D_f(w_i, w_f). \end{aligned} \tag{4}$$

Labor market effects of immigration without work permits

To characterize equilibrium responses to labor supply shocks, we define $\epsilon^D_{\ell_x,w_y}$ as the elasticity of labor demand for worker type x with respect to wages in sector y, and $\epsilon^S_{\ell_x,w_x}$ as the elasticity of salaried labor supply in sector $x \in i, f$. For labor supply shocks R_i and R_f to the informal and formal sectors respectively, the system of market clearing conditions becomes:

$$\begin{aligned} \epsilon^D_{\ell_i,w_i} d\ln w_i + \epsilon^D_{\ell_i,w_f} d\ln w_f &= \epsilon^S_{\ell_i,w_i} d\ln w_i + R_i \\ \epsilon^D_{\ell_f,w_i} d\ln w_i + \epsilon^D_{\ell_f,w_f} d\ln w_f &= \epsilon^S_{\ell_f,w_f} d\ln w_f + R_f \end{aligned} \tag{5}$$

This framework accommodates both a purely informal labor supply shock ($R_i = R > 0$, $R_f = 0$) representing refugee arrivals, and the subsequent effects of work permit provision, which we model as labor reallocation ($R_i = -r/N_i < 0$, $R_f = r/N_f > 0$) where r denotes

[12] For a model where natives can move between formal and informal jobs, see Meghir et al. (2015).

workers transitioning from informal to formal employment, and N_i and N_f denote the number of informal and formal workers at baseline.

To derive closed-form solutions, we specify $\tau(\ell_i) = \ell_i^{1+\gamma}$ with $\gamma > 0$, capturing the convex cost structure of informal hiring. The key labor demand elasticities with respect to informal wages are:

$$\begin{aligned} \epsilon_{L_i,w_i} &= -\frac{1-\rho-(\alpha-\rho)s_f}{(1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f+(1-\rho)s_i]}, \\ \epsilon_{L_f,w_i} &= -\frac{(\alpha-\rho)s_i}{(1-\rho+\gamma)(1-\rho)-(\alpha-\rho)[(1-\rho+\gamma)s_f+(1-\rho)s_i]}, \end{aligned} \tag{6}$$

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

These elasticities reveal two fundamental relationships:

$$\begin{aligned} &\epsilon_{L_i,w_i} < 0 \\ &\epsilon_{L_f,w_i} \begin{cases} > 0 & \text{if } \alpha < \rho \\ < 0 & \text{if } \alpha > \rho \end{cases} \end{aligned} \tag{7}$$

First, $\epsilon_{L_i,w_i} < 0$ for all admissible parameter values, confirming that informal labor demand is downward-sloping. This implies that an informal labor supply shock necessarily reduces informal wages, except in the knife-edge case of perfectly elastic native informal labor supply. Second, the sign of ϵ_{L_f,w_i} depends critically on whether α exceeds ρ, determining whether the immigration shock generates complementarity or substitutability effects in the formal sector.

The economic intuition underlying equation 7 centers on the competing forces of technological complementarity and diminishing returns. The change in the marginal productivity of formal workers when informal employment increases is:

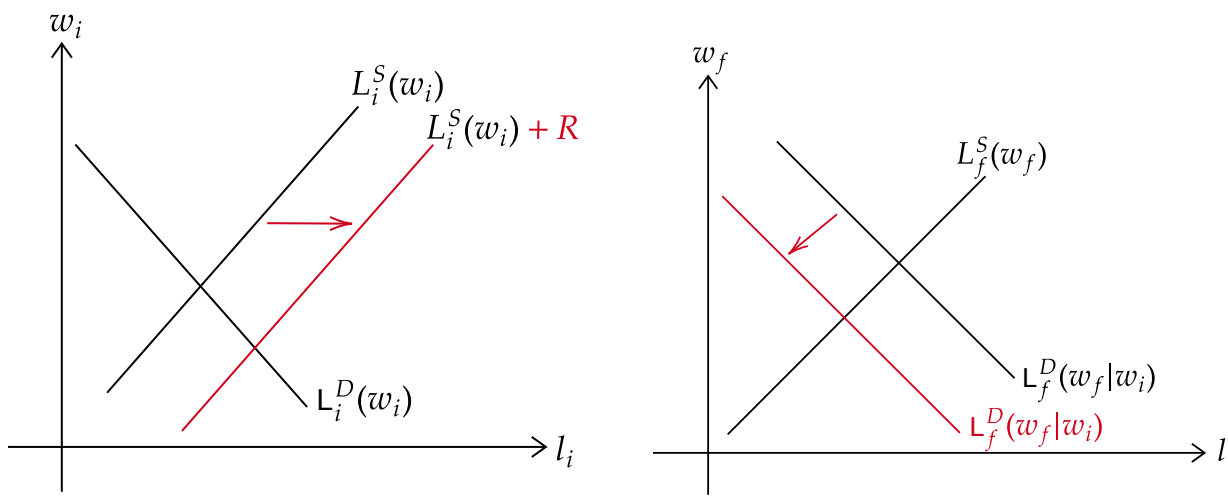
$$\frac{\partial(\ln \partial F/\partial L_f)}{\partial L_i} = (\alpha-\rho)L_i s_i$$

When $\alpha > \rho$—such as under constant returns to scale ($\alpha = 1$) with imperfect substitutability ($\rho < 1$)—hiring additional informal workers enhances formal worker productivity through Q-complementarity, generating $\epsilon_{L_f,w_i} < 0$. As α decreases, the addition of any worker leads to productivity declines for existing workers due to diminishing returns. When $\alpha < \rho$, the productivity losses from diminishing returns dominate the complementarity gains. In this case, an informal labor supply increase that reduces informal wages incentivizes firms to substitute away from formal workers, yielding $\epsilon_{L_f,w_i} > 0$.

The magnitude of the elasticity of substitution $\sigma = \frac{1}{1-\rho}$ proves crucial for determining equilibrium responses. High substitutability between informal and formal workers amplifies firms' incentives to reallocate labor across sectors in response to relative wage changes.

Figure 3 illustrates these dynamics under high substitutability ($\alpha < \rho$). The refugee labor supply shock shifts the informal labor supply curve rightward from $L_i^S(w_i)$ to $L_i^S(w_i) + R$. Panel A demonstrates that informal wages necessarily decline unless labor supply is perfectly elastic. Panel B shows the spillover effects: declining informal wages shift formal labor demand leftward from $L_f^D(w_f|w_i)$ to $L_f^D(w_f|w_i')$, reducing both formal wages and employment as firms substitute toward cheaper informal labor.

Figure 3: Informal LS shock when $\alpha < \rho$



(a) Informal

(b) Formal

Heterogeneous effects across worker groups

The theoretical framework generates heterogeneous predictions across worker subgroups that provide additional empirical tests. First, Colombia's binding minimum wage in the formal sector constrains low-skilled workers' wage adjustment. Consequently, formal labor demand shifts translate entirely into employment changes rather than wage changes for low-skilled natives, implying larger percentage employment effects for this group.

Second, the magnitude of transitions from salaried to non-salaried work depends on the curvature of the earnings function $g(\cdot)$ in non-salaried activities. Appendix Table G.6 documents that this curvature varies systematically across demographic groups, being lower (implying greater responsiveness) for men and declining with education levels. These pat-

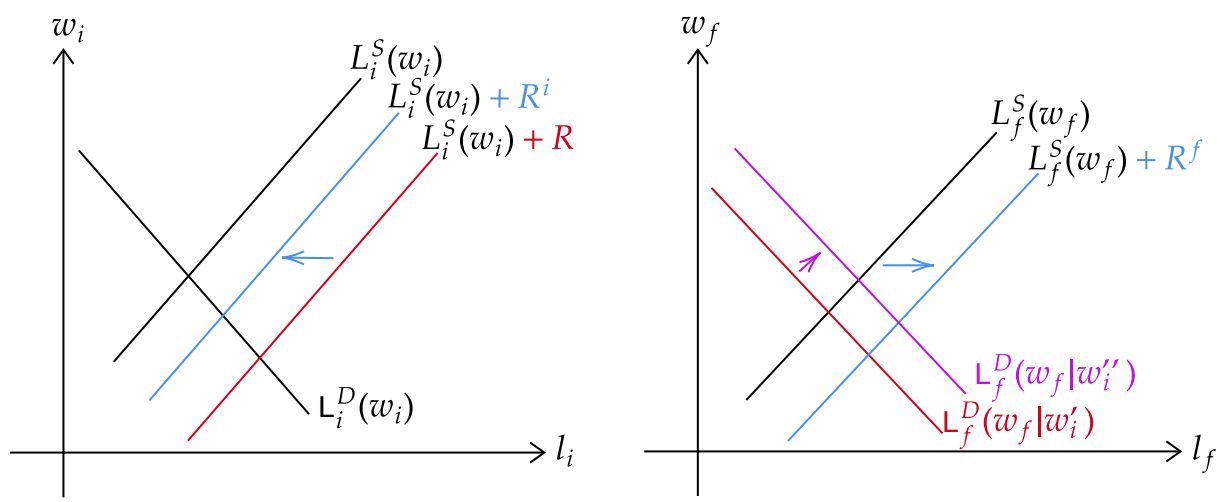
terns predict that low-skilled men should exhibit the strongest transitions to non-salaried employment following informal wage declines.

Work permit effects

Granting work permits to refugees generates two distinct mechanisms, illustrated in Figure 4 under the large substitutability case ($\alpha < \rho$). The primary effect involves sectoral reallocation: permit-holders migrate from informal to formal employment to capture higher formal sector wages. This shifts informal labor supply leftward from $L_i^S(w_i) + R$ to $L_i^S(w_i) + R^i$ while shifting formal labor supply rightward from $L_f^S(w_f)$ to $L_f^S(w_f) + R^f$.

The direct consequences are intuitive: reduced competition in the informal sector benefits native informal workers, while increased competition in the formal sector disadvantages native formal workers. However, a crucial equilibrium effect emerges through wage adjustments. Rising informal wages induced by reduced informal competition cause firms to substitute toward formal workers, shifting formal labor demand rightward from $L_f^D(w_f|w_i')$ to $L_f^D(w_f|w_i'')$. Under high substitutability, this demand increase partially offsets the direct displacement effects, generating a net increase in total formal employment despite displacement of native formal workers.

Figure 4: Effects of Work Permits when $\alpha < \rho$



(a) Informal

(b) Formal

Testable predictions

Our theoretical framework yields six empirically testable predictions regarding the labor market effects of informal immigration and work permit provision:

Effects of informal labor supply shocks:

(1) Native informal salaried employment decreases due to increased sectoral competition.

(2) Native formal salaried employment decreases (increases) when informal and formal labor are substitutes (complements)

(3) Natives transition from informal salaried to non-salaried employment, with stronger effects for low-skilled men

Effects of work permit provision:

(4) Native informal salaried employment increases due to reduced sectoral competition

(5) Native formal salaried employment decreases due to increased sectoral competition

(6) Total formal employment in the economy increases when informal and formal labor are substitutes

Sections 4 and 5 provide empirical tests of predictions (1)–(3) and (4)–(5) respectively. Data limitations prevent direct testing of prediction (6); instead, we employ the calibrated model to quantify this equilibrium effect.

4 The Effects of Informal Labor

In this section, we estimate the causal impact of Venezuelan migration on the Colombian labor market. Our identification strategy employs a shift-share design where the share represents weighted distance-to-border exposure and the shift captures the annual Venezuelan population in Colombia. However, event study estimates reveal systematic pre-trends in our data, indicating that our distance exposure measure correlates with unobserved local labor market trends. This correlation violates the share-exogeneity assumption within standard shift-share designs (Goldsmith-Pinkham et al., 2020). To address this identification challenge, we implement the Synthetic Instrumental Variables (SIV) methodology developed by Gulek and Vives-i Bastida (2023), which relaxes share-exogeneity by using synthetic controls to partial out unmeasured confounders embedded in pre-treatment trends.

4.1 Constructing the distance instrument

We construct our distance exposure measure Z_r following established approaches in the forced migration literature (Angrist and Kugler, 2003). For each Colombian department r, we calculate the population-weighted average inverse distance to the 23 Venezuelan states s:

$$Z_r = \sum_{s=1}^{23} \pi_s \frac{1}{d_{r,s}}, \tag{8}$$

where $d_{r,s}$ represents travel distance between department and state capitals, and π_s denotes the population share of Venezuelan state s. We standardize Z_r to have zero mean and unit variance, ensuring economically interpretable coefficients in our reduced-form specifications.

To evaluate the instrument's predictive power, we estimate an event study specification relating the Venezuelan-to-native ratio $R_{r,t}$ to our distance exposure:

$$R_{r,t} = \sum_{j \neq 2015} \lambda_j (\text{year}_j \times Z_r) + f_r + f_t + \epsilon_{r,t} \tag{9}$$

where year dummies interact with distance exposure, and f_r and f_t represent region and year fixed effects, respectively.

Figure 5 presents the estimated coefficients λ_j from equation 9. The results confirm that distance exposure strongly predicts Venezuelan settlement patterns across all post-treatment periods, with a joint F-statistic of 280 for years 2016–2019. The treatment intensity increases monotonically after 2015, reaching 2 percentage points by 2018 for a one standard deviation increase in distance exposure.[13]

We complete our shift-share instrument by interacting the share Z_r with the shift S_t: $Z_{r,t} = Z_r \times S_t$, where S_t represents the total Venezuelan population in Colombia in year t.

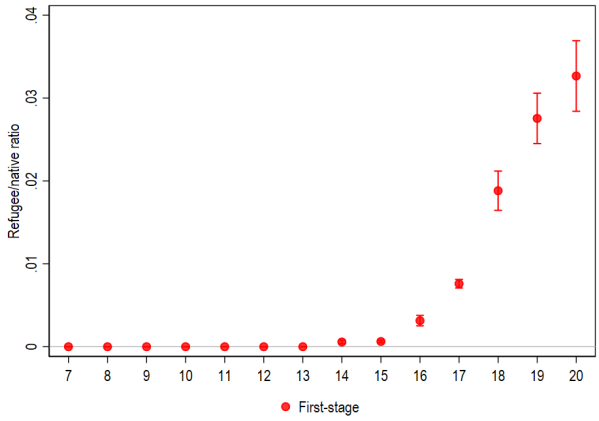
4.2 The pre-trend problem

To visually and flexibly assess the pattern of outcomes $y_{r,t}$ captured by distance exposure relative to the beginning of the refugee crisis, we estimate an event study design. The basic nonparametric event study specification takes the form

$$y_{r,t} = \sum_{j \neq 2015} \theta_j (\text{year}_j \times Z_r) + f_r + f_t + \epsilon_{r,t}, \tag{10}$$

[13] In the GEIH data, we can separate between Venezuelans and Colombians only after 2013. In our analysis, we assume that there were no Venezuelans in Colombia before this date, which is a safe assumption since in 2014 only 0.16% of the working-age population were Venezuelan.

Figure 5: Event Study of the First-stage



Notes: The regression equation is: $R_{i,t} = \sum_{j \neq 2010} \theta_j (\text{year}_j \times Zr) + f_r + f_t + \eta_{r,t}$, where the instrument Z_r is standardized to have mean zero and standard deviation of one to have economically meaningful coefficients, f_r and f_t are region and year fixed effects. Standard errors are clustered at the region level. The 95% confidence intervals are shown.

where the distance share Z_i is interacted with year dummies year_j, and f_i and f_t are region and year fixed effects. The coefficients θ_j compare the outcomes in regions close to and away from the border. Causal interpretation requires that, absent treatment, these regions would follow parallel trajectories. However, our event study estimates reveal systematic pre-treatment differences in trends across multiple labor market outcomes.

Specifically, regions closer to the Venezuelan border experienced declining salaried employment rates relative to interior regions throughout 2007–2015, creating negative trends correlated with our instrument. Visual inspection across various outcomes suggests that no single parametric correction (such as linear time trends) adequately addresses this pervasive pattern. These findings raise fundamental concerns about the validity of distance-based instruments widely employed in the Venezuelan migration literature (Caruso et al., 2021; Delgado-Prieto, 2024; García-Suaza et al., 2024). To make progress, we employ the synthetic IV (SIV) methodology developed by Gulek and Vives-i Bastida (2023).

4.3 The SIV solution

The SIV methodology addresses our identification challenge through a two-step procedure that combines synthetic control (SC) methods with instrumental variables estimation. We provide intuitive exposition here, with technical details in Appendix H.

The procedure operates as follows: First, we solve the standard synthetic control optimization problem to construct weights that best replicate pre-treatment outcomes for treated

units. These weights generate synthetic versions of all variables: outcomes ($\hat{y}_{rt}^{SC}$), treatment ($\hat{R}_{rt}^{SC}$), and instruments ($\hat{Z}_{rt}^{SC}$). Second, we construct "debiased" variables by subtracting synthetic counterparts from observed data: $\tilde{y}_{rt} = y_{rt} - \hat{y}_{rt}^{SC}$, $\tilde{R}_{rt} = R_{rt} - \hat{R}_{rt}^{SC}$, and $\tilde{Z}_{rt} = Z_{rt} - \hat{Z}_{rt}^{SC}$. Finally, we estimate our desired specifications using these debiased variables.

Intuitively, this debiasing procedure addresses confounding by construction: synthetic controls eliminate systematic pre-treatment differences between treated and control units. However, endogenous location choices by migrants remain a concern, which our instrumental variables approach continues to address. Thus, SIV simultaneously handles unobserved confounding through SC and endogenous treatment assignment through IV.

In Section 4.5, we show the reduced-form results by estimating equation 10 separately using IV and SIV, with the only difference being that the instrument and the outcome are their debiased versions in SIV. To show economically meaningful results, we also present 2SLS estimates of the following system of equations using the debiased data:

$$\begin{aligned} \text{LATE:} \quad \tilde{y}_{i,t} &= \beta^{LATE} \tilde{R}_{i,t} + f_i + f_t + \epsilon_{i,t} \\ \tilde{R}_{i,t} &= \theta \tilde{Z}_{i,t} + g_i + g_t + \eta_{i,t} \end{aligned} \tag{11}$$

4.4 Threats to identification

Our empirical strategy compares regions close to the border with those further away. There are three main reasons why this comparison may not be valid.

First, border regions may experience differential impacts from Venezuela's economic crisis through trade relationships rather than migration. However, Venezuela represented a minor trade partner for all Colombian regions prior to the crisis. Appendix Figure C.5 documents that trade flow changes show no systematic correlation with our distance instrument, supporting the exclusion restriction.

Second, the Stable Unit Treatment Value Assumption (SUTVA) could be violated if labor or capital mobility generates spillover effects across regions. Empirically, Appendix Figure C.6 shows border regions experienced modest increases in out-migration and decreases in in-migration, but these effects are economically small and would only marginally bias our estimates toward zero. For capital spillovers, Gulek (2025) demonstrates that immigration-induced trade effects matter primarily for economically developed cities. Since Venezuelan migrants concentrate in less-developed border regions, capital reallocation concerns appear minimal.

Third, as we show later in this section, Venezuelan migration reduces native salaried employment, potentially changing the composition of workers whose wages we observe. Without

individual-level panel data, we cannot fully address this selection bias. Consequently, we focus primarily on employment outcomes, which avoid compositional concerns, while presenting wage results as suggestive evidence in the Appendix Figure C.3.

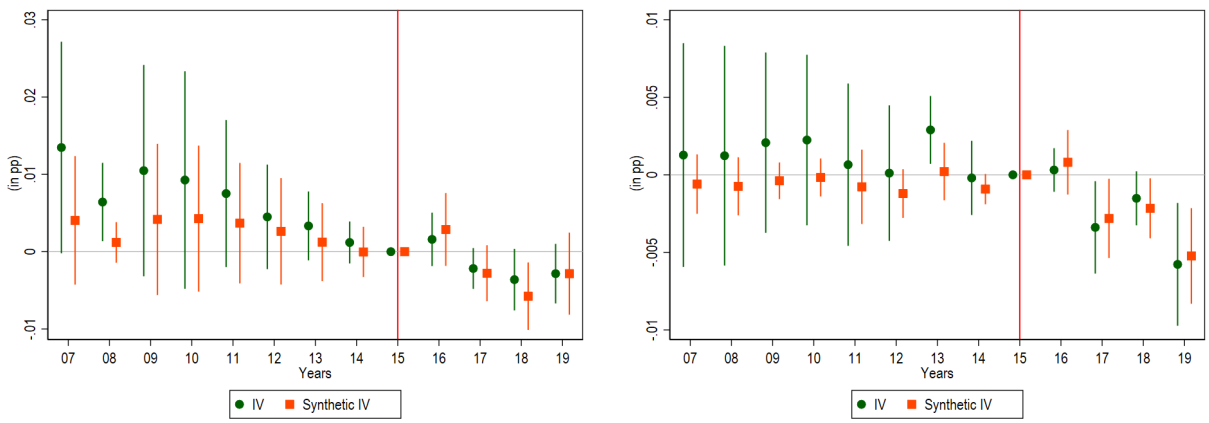
4.5 Empirical Results

Event study estimates

Figure 6 compares event study estimates using standard IV versus SIV methodologies for native salaried employment rates in informal (Panel A) and formal (Panel B) sectors.

Figure 6: Effects on Native Salaried Employment

(a) Informal Employment (b) Formal Employment



Note: IV estimates come from the reduced form regression: $y_{i,t} = \sum_{j \neq 2015} \theta_j (year_j * Z_i) + f_i + f_t + \epsilon_{i,t}$. Synthetic IV estimates come from the same reduced-form design where the dependent variable and the instrument are the debiased versions after applying the first step of the SIV procedure described in text. Standard errors are clustered at the department level.

In Panel A, the IV estimates exhibit a pronounced negative trend between 2007–2014, indicating that border regions experienced systematically declining informal employment rates relative to interior regions well before Venezuelan migration began. This pre-trend likely violates the parallel trends assumption and suggests our instrument correlates with unobserved confounders affecting border region labor markets. The SIV estimates eliminate this problematic pre-trend, confirming effective synthetic control matching in the pre-treatment period (Gulek and Vives-i Bastida, 2023). Post-treatment estimates reveal substantial declines in native informal salaried employment during 2017–2019, with effects becoming increasingly negative over time.

Note that in 2017 and 2018, the SIV estimates are larger in absolute terms than the IV estimates. This means that, absent the migrant treatment, informal salaried employment in

regions close to the border would have risen relative to other regions. Put differently, the negative trend between 2007–2014 not only slows down but also flips sign in the post-period. This is not surprising as the GDP growth rate also flips sign during this period. The GDP per capita in Colombia increased from $4762 in 2007 to $8167 in 2014, but then decreased to $5936 in 2016. If the unobserved confounders leading to the negative trend in the pre-period were related to economic growth, for example, if the more developed nonborder regions spearheaded GDP growth, then the 2015 slowdown could have generated this change in trend.

In Panel B, we observe formal native salaried employment losses in the post-period. Since the instrument does not predict systematic differences between border and non-border regions in the pre-period, relaxing the share exogeneity assumption via SIV does not change our conclusions.

These patterns validate key theoretical predictions. Before work permits became widely available in 2018, Venezuelan migration represented a predominantly informal labor supply shock. The observed decline in native informal salaried employment directly confirms Prediction 1 from our theoretical framework. Simultaneously, the decline in native formal salaried employment supports Prediction 2 under the substitutability case, indicating that informal and formal workers function as gross substitutes rather than complements in production.

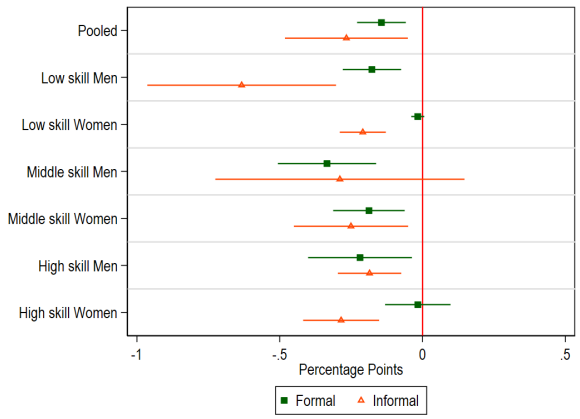
2SLS estimates

Figure 7 presents 2SLS estimates from equation 11 across different worker categories and skill groups. Panel A compares informal versus formal salaried employment effects, while Panel B examines salaried versus non-salaried employment patterns.

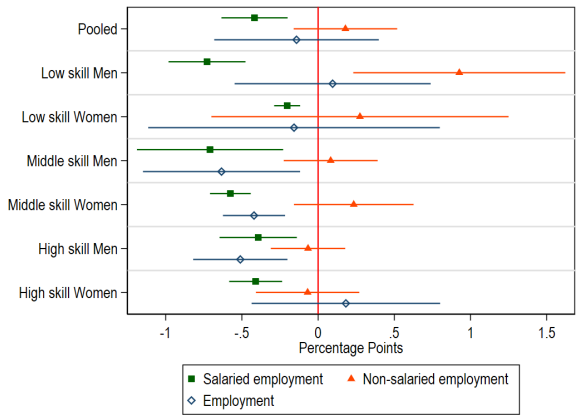
The first row of Figure 7a shows that that the mostly informal immigration shock caused significant declines in the informal and formal native salaried employment rates. A 1 pp increase in migrant/native ratio decreases the native salaried employment rate by 0.27 pp in the informal sector and 0.14 pp in the formal sector. These effects are consistent with model predictions 1&2 and imply that formal and informal labor are highly substitutable in production. To contextualize these magnitudes, consider a municipality with 1,000 working-age natives. Based on 2015 employment patterns, this population would include approximately 270 salaried workers (180 formal, 90 informal). An influx of 10 Venezuelan migrants—representing a 1 percentage point increase in the migrant-to-native ratio—would displace 2.7 informal salaried workers and 1.4 formal salaried workers. Given that Venezuelan migrants maintain a 68% employment rate, this implies that 6.8 new migrant workers displace 4.1 native salaried positions, suggesting substantial displacement.

Figure 7: LATE estimates

(a) Informal vs Formal



(b) Salaried vs Non-salaried



The estimates come from the Synthetic IV methodology described in text. Panel A shows the estimated effects on the informal and formal salaried employment rates across all skill levels, and Panel B shows the estimated effects on salaried and non-salaried employment rates. Standard errors are clustered at the region level.

Rows 2–7 of 7a reveal that employment losses span virtually all education-gender categories, with both informal and formal salaried employment declining across skill groups. This broad-based pattern provides reassurance that our results reflect genuine labor market competition rather than skill-specific demand shocks. Importantly, since only middle- and high-skilled male refugees typically accessed formal employment through work permits, the widespread formal employment losses among all native groups support our theoretical mechanism emphasizing cross-sector substitutability.[14]

Panel B of Figure 7 provides crucial evidence supporting Prediction 3 regarding native transitions to non-salaried work. While salaried employment declines significantly across most skill groups, non-salaried employment increases, particularly among low-skilled workers. Although the non-salaried employment increases lack statistical significance in the aggregate, they are economically meaningful and concentrated among demographic groups predicted by our theoretical framework.

This finding carries important implications for understanding immigration's labor market effects in developing countries where non-salaried positions constitute important margins of the labor market. The offsetting movements in salaried and non-salaried employment mean

[14] Appendix Figure C.2 presents the estimated effects in percentages. It shows that low-skill men lose the largest amount of formal salaried jobs while high-skill natives lose the least. This is consistent with a binding minimum wage for the low-skill market causing the demand adjustments to be more on the employment margin rather than the wage margin.

that total employment rates show minimal change, potentially masking substantial underlying adjustments. Researchers focusing exclusively on aggregate employment rates might erroneously conclude that massive immigration flows generated no labor market impacts. The decomposition reveals that, in Colombia, the relevant outside option to salaried employment is non-salaried work rather than non-employment, especially for low-skilled men.

4.6 Robustness Checks

Industry heterogeneity

To address concerns that our results reflect contemporaneous industry-specific demand shocks rather than immigration effects, we estimate equation 11 using employment rates across major industry categories. Figure C.4 shows the results. Salaried employment rates drop across various industries, except for Health and Social Services. The native escape to non-salaried employment occurs mostly in Manufacturing, Transportation, and to some degree in Agriculture, Health and Social Services, and Education. This transition to non-salaried employment masks employment losses in these industries. The fact that saleried employment decreases in most industries provides evidence against contemporaneous demand shocks.

Quality checks of SIV

The reliability of our results depends critically on SIV's performance. Following Gulek and Vives-i Bastida (2023), we implement three diagnostic checks to validate our methodology:

Strong Debiased First Stage: Since SIV constructs outcome-specific synthetic controls, each specification generates unique debiased instruments and treatments. For SIV to provide valid estimates, the first-stage should remain strong after the debiasing step. Figure H.12 reports F-statistics across our 35 main specifications. The minimum F-statistic exceeds 32, with over 75% of specifications achieving F-statistics above 100, confirming robust instrument strength throughout our analysis.

Effective Pre-Treatment Matching: SIV requires a good fit in the training set. In our case, SIV successfully eliminates pre-trends across our main outcomes, indicating high-quality synthetic control fits. Nonetheless, we further enhance matching performance by incorporating constant shifts (Doudchenko and Imbens, 2016), equivalent to computing synthetic controls using pre-treatment demeaned data (Abadie and Vives-i Bastida, 2022). Figures H.13, H.14, H.15, and H.16 demonstrate that our main conclusions remain robust to this refinement.

Back-Testing for Overfitting: We validate our approach by restricting the training

period from 2007–2015 to 2007–2012, testing whether shorter pre-treatment windows generate similar results. Figure H.17 confirms that our estimates remain stable, indicating that our original eight-year training period does not suffer from overfitting concerns.

4.7 Estimating the elasticity of substitution between informal and formal workers

Our theoretical framework suggested that the effect of undocumented migrants on formal employment would be theoretically indeterminate. The empirical findings demonstrate that Venezuelan migration led to job displacement among natives in both informal and formal sectors, suggesting substantial substitutability between these labor categories. Measuring this degree of substitutability, captured by the parameter ρ, represents a key contribution with broader implications for our subsequent structural analysis. We provide a concise overview of how our empirical estimates combined with data moments allows us to recover the model's structural parameters, with comprehensive estimation details presented in Appendix Section E.

Within our theoretical framework, labor represents the single production input available to firms for expanding output. Therefore, firm size differences reflect solely variations in underlying productivity levels. Higher-productivity enterprises select larger employment levels accordingly. The cost parameter γ, representing the marginal expense of hiring informal labor, depends crucially on the degree to which more productive firms choose formalization along the intensive margin. Across all types of firms, the share parameter η captures the relative efficiency difference between formal and informal employment, making it identifiable through the aggregate informal employment rate observed in the data. The substitution elasticity that governs formal-informal worker relationships derives primarily from the underlying demand-side elasticities.

Our identification strategy yields an estimate of $\rho \approx 0.91$, corresponding to a substitution elasticity $\sigma = \frac{1}{1-\rho}$ of approximately 11. This represents only the second study to quantify this particular elasticity following Gulek (2023), and constitutes the first such estimate for the Colombian labor market. This substantial elasticity estimate aligns well with Colombia's institutional context, where informal work frequently occurs within identical industries and sometimes within the same enterprises as formal employment. Our findings provide empirical support for the perfect substitutability assumption commonly employed in recent structural models of informal labor markets (Ulyssea, 2018, 2020).

5 The Effects of Work Permits

Having established that Venezuelan migration generated substantial labor market displacement through informal labor supply increases, we now examine how work permit policies affect these dynamics. This analysis addresses a fundamental policy question: do work permits mitigate or exacerbate the negative effects of forced migration on native workers? Our theoretical framework predicts that work permits should generate opposing effects in informal and formal sectors—reducing competition in informal markets while increasing it in formal markets—with potentially positive net effects on total formal employment when workers are highly substitutable.

Testing these predictions empirically presents significant challenges. The spatial distribution of the work permit strongly correlates with migration pattern, making it difficult to separate the effects of the work permit from the ongoing impacts of immigration. We address this identification challenge by exploiting institutional features of Colombia's work permit system that created differential treatment intensity across demographic groups, enabling a triple difference-in-differences design that partials out immigration effects.

5.1 Identification

The primary identification challenge stems from the spatial correlation between migration and work permit treatments. Appendix Figure B.1 demonstrates that regions receiving more Venezuelan migrants also received more work permits proportionally. When controlling for migration intensity, insufficient spatial variation remains to identify work permit effects using standard approaches.

This correlation arises because work permits (PEPs) were distributed based on registration at migration offices, which Venezuelan migrants accessed differentially depending on their settlement patterns. Border regions, which received more migrants, naturally also processed more work permits, creating the confounding relationship that threatens identification.

To make progress, we exploit the fact that work permits had little to no bite for women and low-skill men. As shown in Section 2.1, PEP only enabled high-skill and middle-skill men to find formal jobs. Other Venezuelan refugees were not affected. Moreover, the high and middle skill men were impacted by the work permits at different times: while PEP-1 and PEP-2 enabled high-skill venezuelans to find formal jobs starting from the 3rd quarter of 2017, PEP-3 enabled middle-skill men to find formal jobs starting from 2018 q4. This variation across skill cells provides a third dimension in which the work permit treatment intensity varies in addition to space and time. It allows for a triple DiD design that identifies

the effect of work permits by partialing out the effects of immigration.

The triple differences in differences design

Our identification strategy rests on the premise that middle- and high-skilled native men face greater exposure to work permit effects than other demographic groups, since they are closest substitutes to the skilled migrant men who could utilize work permits effectively. We formalize this logic in a triple difference-in-differences framework that compares outcomes across skill groups, time periods, and regions.

We define treatment exposure as:

$$D_{brt} = \begin{cases} 1 & \text{if } b = \text{high-skill men } \& \ t > 2017q3 \ \& \ \text{region} = \text{close to border} \\ 1 & \text{if } b = \text{middle-skill men } \& \ t > 2018q4 \ \& \ \text{region} = \text{close to border} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where treatment switches on when work permits become available to each skill group in border-proximate regions. We use a binary treatment indicator for interpretational clarity, though results remain robust to using the continuous distance exposure measure Z_r (shown in the Appendix).

Formally, we estimate the following regression equation:

$$Y_{brt} = \beta D_{brt} + f_{br} + f_{bt} + f_{rt} + \epsilon_{brt}, \tag{13}$$

where Y_{brt} represents labor market outcomes for demographic group b in region r at time t. The fixed effects structure is crucial for identification: f_{br} (group-region effects) control for time-invariant differences in how regions affect specific demographic groups; f_{bt} (group-time effects) capture economy-wide trends affecting particular skill-gender categories; and f_{rt} (region-time effects) absorb all regional shocks, including ongoing migration impacts.

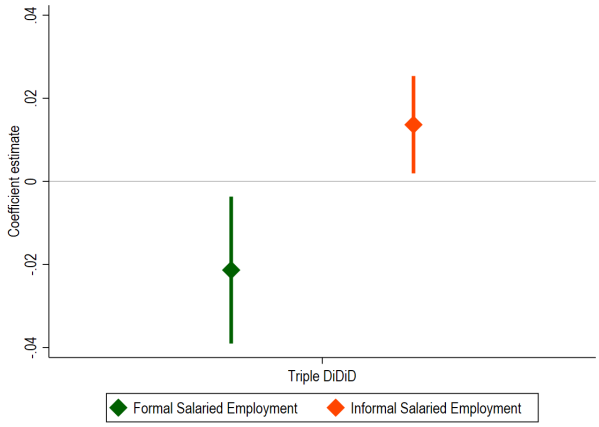
5.2 Empirical results

Main effects on salaried employment

Figure 8 presents our central findings from estimating equation 13 for informal and formal salaried employment rates. The results provide strong support for Predictions 4 and 5 from our theoretical framework.

Work permits generate statistically significant decreases in native formal salaried employment and increases in native informal salaried employment. These intent-to-treat effects align

Figure 8: Triple DiD Estimate



Notes: The figure shows estimated coefficients from equation 13 where the dependent variables are informal and formal salaried employment rates. Treatment exposure is defined as a binary indicator for high-skilled men after 2017Q3 and middle-skilled men after 2018Q4 in border regions. Standard errors are clustered at the skill-group-region level. Confidence intervals show 95% confidence levels.

precisely with theoretical predictions: as work permits enable skilled migrants to transition from informal to formal sectors, they reduce competition in informal markets (benefiting native informal workers) while increasing competition in formal markets (displacing native formal workers).

5.3 Robustness Checks

Our results also remain robust to expanding the pre-period observations (Figure D.7a), using alternative cutoffs (Figure D.7b) or the continuous distance exposure Z_i (Figure D.7c). We also provide the implied event-study estimates of the triple DiD design (Figure D.8) to show that our results are not driven by pre-trends.

We conduct two more exercises to ensure that these effects are not driven by outlier cells within the region-time-skill dimension and report them in Appendix C. First, we estimate region-specific DiD estimates comparing the employment outcomes of high-skill and middle-skill men to the rest, before and after the issuance of work permits. We plot these estimates along the distance exposure dimension in Figure D.9. This figure shows that our results are not driven by an outlier region having skill-cell-specific shocks. Second, we estimate education-sex cell-specific DiDs, comparing employment outcomes in regions close to and away from the border, before and after the issuance of work permits. We plot these estimates for each education-sex cell in Figure D.10. This exercise shows that our results are not driven by an outlier skill cell in one of the control groups. Together, these exercises reveal that our

triple DiD estimates are not driven by outlier cells, which increases the reliability of our conclusions.

5.4 Data limitations and model-based quantification

While we successfully test five of our six theoretical predictions empirically, data limitations prevent direct evaluation of Prediction 6: that work permits increase total formal employment in the economy. The GEIH survey lacks sufficient precision to measure Venezuelan employment across region-skill-time cells, while administrative records of PEP holders cannot be disaggregated by skill categories.

However, having validated the model's other predictions gives us confidence in using the theoretical framework to quantify this final mechanism. This approach follows the broader literature on structural policy evaluation, where validated models provide insights beyond what reduced-form methods can deliver.

How many formal jobs have been created due to work permits?

Given the model parameters, we can estimate the effect of providing work permits to refugees on formal job creation. Here is how we proceed. From Figure 1c, we estimate how many Venezuelans were able to find formal jobs thanks to work permits. This enables us to estimate how much the total informal labor supply in the economy changes. From the model parameters, we determine the informal labor demand elasticity ϵ_{L_i,w_i}. This allows us to determine how much informal wages would change given the change in informal labor supply due to work permits. Lastly, the formal labor demand elasticity times the estimated change in informal wages gives us the percentage change in the formal labor demand, which in turn gives us the number of formal jobs that are created in the economy. We estimate that from 240,000 work permits issued between PEP1–3, around 68,000 Venezuelans became able to transition from informal to formal sectors.[15] This in turn has created around 24,440 formal jobs in the economy. Assuming these jobs paid the mean wage in the formal sector (despite Venezuelans making this transition being more educated than the average formal Colombian worker), the total tax revenue from granting work permits would be around 43 million USD annually.[16]

[15] We also validate this number using admin records. By the end of 2019, 73360 Venezuelans with PEPs were paying social security through their jobs. Using this number as opposed to 68,000 does not change our counterfactual estimates in a meaningful way.

[16] A similar counterfactual exercise can be found in Gulek (2023), who estimates what would have happened if Syrian refugees were all given work permits in Turkey. Since he does not know what ratio of refugees would be able to find formal jobs given work permits, he can only give a range of estimates on the effect of work permits. In contrast, because we know the exact number, we can point estimate the effect of granting

5.5 Do work permits reduce mismatch in the economy?

Our analysis thus far demonstrates that work permits cause sectoral reallocation between informal and formal employment. However, permits may also generate productivity gains by reducing skill mismatch in the economy. Consider a Venezuelan engineer who works informally in a restaurant due to lack of work authorization. Work permits may enable him to transition to formal employment in manufacturing or professional services that better utilize his skills. Such transitions could increase competition among natives in skill-intensive industries while simultaneously generating positive knowledge spillovers that enhance native productivity.

We investigate whether work permits generate such productivity spillovers by examining wage effects across industries that vary in their exposure to skilled migrant workers. Our approach builds on the theoretical framework developed in Appendix Section F, which shows that absent productivity spillovers, the transition of skilled migrants into skill-intensive industries necessarily reduces native wages in those sectors. However, if skilled migrants generate skill-enhancing knowledge spillovers, native wages in skill-intensive industries need not decline and may actually increase.

This framework yields a testable prediction: if work permits generate productivity spillovers, the wage gap between skilled and less-skilled natives should not decrease, and may even increase, in industries with greater exposure to skilled migrant workers. Testing this prediction requires comparing wage effects across industries that differ in their likelihood of employing skilled workers with permits.

Empirical strategy

We employ a quadruple DiD design that extends our triple DiD framework by adding variation across industries. The identification strategy compares wages of more exposed natives (high- and middle-skilled men) to less exposed natives, in regions close to and away from the border, before and after work permits are issued, and across industries that differ in their exposure to skilled migrant workers.

Our estimating equation becomes:

$$ln(w_{ibrt}) = \beta D_{brt} * Exposure_i + f_{ibr} + f_{ibt} + f_{irt} + f_{brt} + \epsilon_{ibrt} \tag{14}$$

where w_{ibrt} represents native wages in industry i, skill group b, region r, at time t. The treatment intensity D_{brt} equals the distance instrument Z_r for middle-skilled men after 2018Q4

work permits on formal job creation.

and for high-skilled men after 2017Q3, and zero otherwise. Industry exposure Exposure_i measures either the share of formal workers or the share of college graduates in industry i, capturing industries' differential capacity to employ skilled migrants with work permits.

The comprehensive fixed effects structure is crucial for identification. Three-way fixed effects f_{ibr} control for time-invariant differences across industry-skill-region combinations, f_{ibt} account for industry-specific skill wage gaps that may evolve over time, f_{irt} capture industry-region-specific trends, and f_{brt} absorb all skill-region-time variation (including the main work permit effects estimated in Section 5). This saturation ensures that we exploit only the differential timing of work permit effects across industries with varying exposure to skilled migrants.

Visual evidence

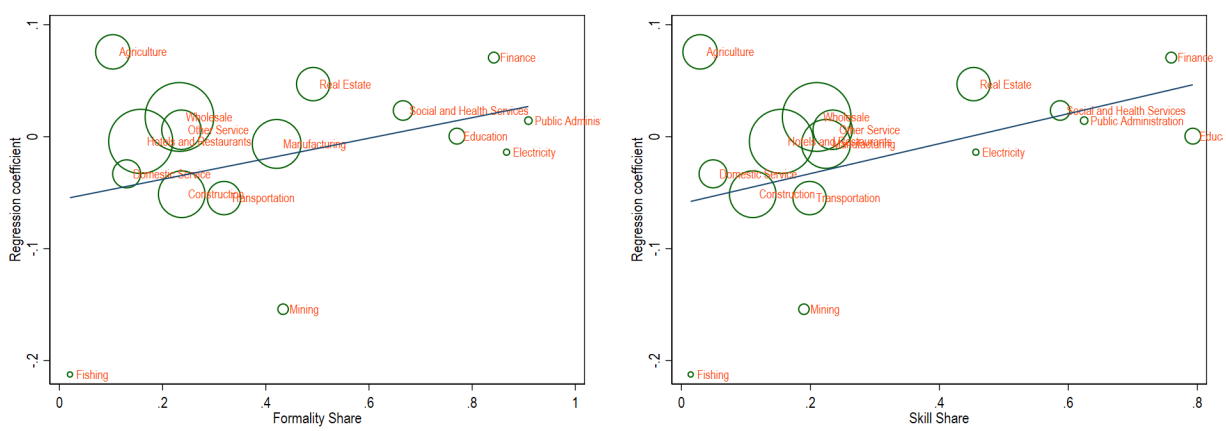
Before presenting formal results, we provide visual evidence by estimating industry-specific triple DiD regressions:

$$ln(w^i_{brt}) = \beta^i D_{brt} + f^i_{br} + f^i_{bt} + f^i_{rt} + \epsilon^i_{brt}, \tag{15}$$

for each major industry i separately. Figure 9 plots the industry specific β^i estimates against formality rate in Panel A and skill-intensity in Panel B.

Figure 9: Visual Quadruple DiD design

(a) Formality (b) Skill



Note: Each circle represents a triple DiD coefficient from equation 15 where the dependent variable is mean native wages. Circle sizes reflect industry employment shares. Panel A plots coefficients against industry formality rates; Panel B against the share of college graduates. The fitted lines represent visual representations of the quadruple DiD design.

The results reveal a striking pattern. In industries with low shares of skilled or formal

workers, such as Agriculture and Construction, skilled natives in border regions experience wage declines following work permit introduction, consistent with increased competition from migrants transitioning to formal employment. However, the industry-specific estimates become progressively less negative, and eventually positive, as we move toward more skill- and formal-intensive industries like Finance and Professional Services.

The positive slopes of the fitted lines provide visual evidence of productivity spillovers. If work permits only generated competition effects, we would expect uniform wage declines across all industries. Instead, the differential pattern suggests that in skill-intensive industries, positive productivity effects offset, and potentially dominate, the negative competition effects.

Formal results

Table 1 presents formal estimates of equation 14 using both formality and skill intensity as exposure measures. We estimate effects across the wage distribution to assess whether results reflect compositional changes versus genuine productivity shifts.

Table 1: Quadruple DiD results

Hourly Earnings	(1) Mean	(2) p10	(3) p25	(4) p50	(5) p75	(6) p90
Formality	0.0457*	0.0851***	0.0629***	0.0405	0.0655**	0.0288
	(0.0272)	(0.0252)	(0.0238)	(0.0310)	(0.0297)	(0.0404)
Skill intensity	0.0854**	0.0947***	0.0907***	0.0818**	0.117***	0.0741
	(0.0329)	(0.0332)	(0.0290)	(0.0402)	(0.0341)	(0.0486)

Note: Estimates from equation 14 where the dependent variable captures native wage statistics at the industry-skill-region-time level. Column 1 shows the effects on mean wages, and columns 2–6 show the effects on the 10th, 25th, 50th, 75th, and 90th percentiles of the wage distribution, respectively. Formality exposure equals the share of formal workers in each industry; Skill intensity equals the share of college graduates. Treatment intensity D_{brt} switches on for high-skilled men after 2017Q3 and middle-skilled men after 2018Q4 in border regions. All specifications include three-way fixed effects. Standard errors clustered at skill-region level. *p<0.10, **p<0.05, ***p<0.01.

The formal results confirm the visual evidence. Column 1 shows that average wages increase significantly in industries with greater exposure to skilled migrants: a 10 percentage point increase in industry formality is associated with 0.46% higher wages for skilled natives in border regions, while a similar increase in skill intensity yields 0.85% higher wages.

To interpret these magnitudes, consider the difference between Finance (90% formal, 45% college-educated) and Agriculture (30% formal, 5% college-educated). Our estimates suggest

that skilled natives in border regions experience wages that are 2.7% higher in Finance than Agriculture after the introduction of the work permit using formal exposure or 3.4% higher using the intensity of the skill. These represent economically meaningful spillover effects.

Crucially, the positive effects persist across most of the wage distribution (columns 2–6), which rules out compositional explanations. If work permits simply enabled migrants to displace the lowest-earning natives in skill-intensive industries, we would observe positive effects only at higher percentiles. Instead, we find significant positive effects at the 10th, 25th, and 75th percentiles for both exposure measures, suggesting genuine productivity enhancements rather than changes in worker composition.

Discussion

These findings provide suggestive evidence that work permits generate skill-enhancing productivity spillovers by enabling better matching between workers and jobs. When skilled migrants can work formally in industries that utilize their human capital, they can enhance the productivity of native co-workers, for example, through knowledge transfer or complementary skills.

However, we acknowledge several limitations in interpreting these results as definitive evidence of productivity spillovers from reduced mismatch. First, while our comprehensive fixed effects structure controls for numerous confounding factors, we cannot entirely rule out alternative explanations such as industry-specific demand shocks that coincide with work permit timing. Second, our design captures the reduced-form effect of improved skill matching rather than isolating specific spillover mechanisms. Third, the simultaneous occurrence of migration and work permit shocks limits our ability to separate these effects completely.

Despite these limitations, several factors strengthen confidence in the spillover interpretation. The systematic pattern across industries with different exposure levels is difficult to explain through alternative mechanisms. The persistence of effects across the wage distribution rules out simple compositional explanations. Most importantly, the direction of effects, positive spillovers in skill-intensive industries where migrants with permits are most likely to work, aligns precisely with theoretical predictions.

These findings have important policy implications beyond the job creation effects documented earlier. If work permits generate productivity spillovers by reducing skill mismatch, the economic benefits extend beyond the fiscal revenues from increased formal employment. Enhanced productivity in skill-intensive industries could generate broader economic growth that benefits both migrants and natives.

Moreover, the spillover effects provide an additional economic rationale for work permit

policies. While our earlier analysis showed that permits create 24,440 formal jobs, the productivity gains documented here represent an additional, unmeasured benefit stream. Future research investigating occupational transitions and firm-level productivity changes could help quantify these spillover effects more precisely.

6 Conclusion

This paper provides the first comprehensive analysis of how informal immigration and work permit policies jointly affect labor markets in developing economies. Using the Venezuelan refugee crisis in Colombia as a quasi-experiment, we develop a theoretical framework that incorporates both formal-informal and salaried-nonsalaried employment distinctions, and implement novel empirical strategies to separately identify immigration and policy effects.

Our findings demonstrate that predominantly informal immigration generates displacement across both informal and formal sectors, with displaced natives transitioning to nonsalaried employment rather than unemployment. These results indicate a high substitutability between informal and formal workers (elasticity ≈ 11) and highlight the importance of non-salaried employment as an adjustment mechanism in the labor markets of developing countries, dynamics that would be missed by focusing solely on aggregate employment effects.

Work permits fundamentally alter these dynamics by enabling sectoral reallocation that generates net economic benefits. While permits reduce informal sector competition and increase formal sector competition, the resulting wage changes induce firms to substitute toward formal employment. In our setting, this creates approximately 24,440 additional formal jobs beyond direct transitions. We also find suggestive evidence that permits generate productivity spillovers by reducing skill mismatch.

These findings contribute to several literatures. For the economics of immigration, we demonstrate that the effects of migrants with and without work authorization differ substantially, and that standard approaches may underestimate displacement by overlooking nonsalaried employment transitions. For research on informality, we provide the first estimate of the elasticity of substitution between informal and formal workers in Colombia and show how immigration can affect formalization incentives. For policy analysis, we document the distributional effects of work permits and provide suggestive evidence of productivity spillovers from reduced mismatch.

The results have important policy implications as refugee populations continue growing globally while most lack work authorization. Our analysis suggests that work permits not only benefit refugees but can generate substantial economic returns for host countries

through increased formalization, job creation, and productivity gains. As developing countries host increasing numbers of displaced populations, understanding how policy design shapes integration outcomes becomes crucial for both humanitarian and economic reasons.

Work permits represent only one policy tool available to policymakers managing immigration shocks. Access to education, healthcare, language training, and credential recognition may have equally important consequences for long-run labor market integration. Future theoretical and empirical work can build on our framework to separate these forces and understand how different policy combinations affect both short-run displacement and long-run economic outcomes.

References

Abadie, Alberto and Jaume Vives i Bastida, "Synthetic controls in action," *arXiv preprint arXiv:2203.06279*, 2022.

Amaral, Pedro S and Erwan Quintin, "A competitive model of the informal sector," *Journal of monetary Economics*, 2006, *53* (7), 1541–1553.

Andrade, Gustavo Henrique De, Miriam Bruhn, and David McKenzie, "A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms," *The World Bank Economic Review*, 2016, *30* (1), 24–54.

Angrist, Joshua D and Adriana D Kugler, "Protective or counter-productive? labour market institutions and the effect of immigration one unatives," *The Economic Journal*, 2003, *113* (488), F302–F331.

Bahar, Dany, Ana María Ibáñez, and Sandra V Rozo, "Give me your tired and your poor: Impact of a large-scale amnesty program for undocumented refugees," *Journal of Development Economics*, 2021, *151*, 102652.

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

Bosch, Mariano and Julen Esteban-Pretel, "Job creation and job destruction in the presence of informal markets," *Journal of Development Economics*, 2012, *98* (2), 270–286.

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

_ , "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.

Caruso, German, Christian Gomez Canon, and Valerie Mueller, "Spillover effects of the Venezuelan crisis: migration impacts in Colombia," *Oxford Economic Papers*, 2021, *73* (2), 771–795.

Clemens, Michael, Cindy Huang, Jimmy Graham et al., "The economic and fiscal effects of granting refugees formal labor market access," *Center for Global Development Working Paper*, 2018, *496.*

Delgado-Prieto, Lukas, "Immigration, Wages, and Employment under Informal Labor Markets," 2024.

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

Elias, Ferran, Joan Monras, and Javier Vázquez-Grenno, "Understanding the effects of granting work permits to undocumented immigrants," *Journal of Labor Economics*, 2025, *43* (3), 000–000.

European Commission, "EU Solidarity with Ukraine," https://eu-solidarity-ukraine.ec.europa.eu/information-people-fleeing-war-ukraine/fleeing-ukraine-your-rights-eu_en. Accessed: 2023-03-10.

García-Suaza, Andrés, Juan Miguel Gallego Acevedo, Juan Diego Mayorga, Angie Mondragón-Mayo, Carlos Sepúlveda, and Alexander Sarango-Iturralde, "Venezuelan Immigration in Colombia: Labor Market Dynamics," *The Center for Growth and Opportunity*, 2024.

Glitz, Albrecht, "The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in Germany," *Journal of Labor Economics*, 2012, *30* (1), 175–213.

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

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

Gulek, Ahmet, "Formal Effects of Informal Labor: Evidence from the Syrian Refugees in Turkey," 2023. Available at https://ahmetgulek.github.io/FEoIL.pdf.

_ , "Effects of Immigrants on Non-host Regions Evidence from the Syrian Refugees in Turkey," 2025. Available at https://ahmetgulek.github.io/Gulek_jmp.pdf.

_ **and Jaume Vives i Bastida**, "Synthetic Synthetic IV estimation in panels," 2023. Available at https://jvivesb.github.io/jmp/draft_siv.pdf.

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

Lebow, Jeremy, "The labor market effects of Venezuelan migration to Colombia: reconciling conflicting results," *IZA Journal of Development and Migration*, 2022, *13* (1).

_ , "Immigration and occupational downgrading in Colombia," *Journal of Development Economics*, 2024, *166*, 103164.

Llull, Joan, "The effect of immigration on wages: exploiting exogenous variation at the national level," *Journal of Human Resources*, 2018, *53* (3), 608–662.

Meghir, Costas, Renata Narita, and Jean-Marc Robin, "Wages and informality in developing countries," *American Economic Review*, 2015, *105* (4), 1509–46.

Monteiro, Joana CM and Juliano J Assunção, "Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises," *Journal of Development economics*, 2012, *99* (1), 105–115.

Paula, Aureo De and José A Scheinkman, "The informal sector: An equilibrium model and some empirical evidence from Brazil," *Review of Income and Wealth*, 2011, *57*, S8–S26.

Penaloza-Pacheco, Leonardo, "Living with the neighbors: the effect of Venezuelan forced migration on the labor market in Colombia," *Journal for Labour Market Research*, 2022, *56* (1), 1–32.

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.

Rauch, James E, "Modelling the informal sector formally," *Journal of development Economics*, 1991, *35* (1), 33–47.

Rocha, Rudi, Gabriel Ulyssea, and Laísa Rachter, "Do lower taxes reduce informality? Evidence from Brazil," *Journal of development economics*, 2018, *134*, 28–49.

Santamaria, Julieth, "when a stranger shall sojourn with thee': The impact of the venezuelan exodus on colombian labor markets," *University of Minnesota, Department of Applied Economics, working paper*, 2020, *51422.*

Schramm, Heidi R, "The equilibrium effects of income taxation on formal and informal labor markets," in "NEUDC 2014 Working Paper. Presented at the Northeast Universities Development Consortium Conference" Boston University Boston 2014.

Tumen, Semih, "The economic impact of Syrian refugees on host countries: Quasi-experimental evidence from Turkey," *American Economic Review*, 2016, *106* (5), 456–60.

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.

Appendix

A Data Appendix

Table A.1: Summary Statistics

Sex	Pooled		Men			Women	
Skill	Pooled	Low	Medium	High	Low	Medium	High
Panel A: Employment							
All	0.643	0.786	0.775	0.759	0.439	0.523	0.664
Formal	0.213	0.131	0.321	0.491	0.047	0.160	0.437
Informal	0.430	0.655	0.455	0.267	0.392	0.363	0.227
Panel B: Salaried Employment							
All	0.242	0.227	0.384	0.379	0.087	0.221	0.347
Formal	0.151	0.100	0.255	0.313	0.031	0.128	0.280
Informal	0.091	0.127	0.129	0.066	0.056	0.094	0.067
Panel C: Non-salaried Employment							
All	0.372	0.555	0.359	0.284	0.349	0.289	0.219
Formal	0.035	0.028	0.036	0.086	0.014	0.021	0.064
Informal	0.337	0.527	0.323	0.197	0.335	0.268	0.155
Panel D: Income (in thousands)							
All	807	565	778	1772	363	523	1305
Formal	1298	807	927	2067	649	724	1570
Informal	538	512	663	1184	324	422	743
Panel E: Wages (Hourly)							
All	4661	2905	3957	10223	2514	3321	8286
Formal	7101	3841	4451	11449	3421	3904	9341
Informal	3330	2701	3576	7773	2389	3032	6043
Panel F: Wages (salaried empoyment)							
All	4926	3208	3739	9109	2831	3243	7043
Formal	5949	3668	4036	9876	3219	3571	7718
Informal	3169	2823	3121	5455	2607	2781	4224

Note: Sample is restricted to those between ages of 15–65 in years 2007–2015. Source: Author's calculations using GEIH.

Table A.2: Types of Employment according to GEIH

Sex	Pooled	Men			Women		
Skill	Pooled	Low	Medium	High	Low	Medium	High
Works in a private company	0.355	0.276	0.459	0.442	0.170	0.390	0.478
Government	0.061	0.006	0.051	0.149	0.006	0.027	0.171
Domestic	0.041	0.003	0.003	0.001	0.166	0.073	0.015
Self-employed	0.447	0.582	0.409	0.323	0.555	0.420	0.277
Employer	0.041	0.052	0.047	0.066	0.020	0.025	0.029
Unpaid worker in own home	0.035	0.024	0.021	0.016	0.072	0.058	0.027
Unpaid worker at another home	0.003	0.002	0.002	0.002	0.006	0.005	0.003
Day worker	0.016	0.054	0.009	0.001	0.005	0.001	0.000
Other	0.001	0.001	0.001	0.001	0.001	0.002	0.001

Note: Sample is restricted to those between ages of 15–65 in years 2007–2015. Salaried job is defined as working in a private company. Non-salaried jobs are all types of jobs other than salaried jobs and government jobs. Source: Author's calculations using GEIH.

B Details on Work Permits

The *Permiso Especial de Permanencia* (PEP) was a work permit created by Colombia in 2017 specifically to address the large influx of Venezuelan migrants. It was unprecedented in its scope and ease of access and generosity. As will be detailed below, there were few if any requirements, and the application was completed online. Once approved (usually in under two weeks), recipients then simply had to print out their ID number. The permits were granted for 90 days but automatically renewed for a period of 2 years unless an infraction took place. Holders of a PEP could work legally in any capacity, but also access the public healthcare and educational system. For this reason, they were attractive for all ages.

On July 25th of 2017, the ministry of foreign affairs created the first PEPs under resolution 5797[17]. They were implemented by the office of migration *(Unidad Administrativa Especial Migración Colombia)* under resolution 1272 three days later. During this first wave of PEP, applications were received between August 3rd and October 31st 2017. Eligibility was restricted to Venezuelans who had entered the country legally before July 28th and did not have a criminal record. This is an important point. To enter Colombia legally, migrants needed to have passed through an approved port of entry and received a tourist stamp in their passport. Venezuelan passports are notoriously expensive, with some estimates putting the cost upwards of 200$.[18] This implies that most Venezuelans who hold a passport are middle/high skill.

Resolutions 0740 and 0361, published on February 5th and 6th of 2018, implemented a second round of PEPs. These were available to Venezuelan migrants who, as before, had entered the country legally before February 2nd. The application window was open between February 7th and June 7th 2018.

In March of 2018, the Colombian government set out to register undocumented Venezuelans for the stated purpose of designing better public policies (Decree 542). They called it RAMV (*Registro Administrativo de Migrantes Venezolanos*). This registry was open between April 6th and June 8th of 2018. During this period, surveyors travelled to areas with high concentrations of undocumented Venezuelans and carried out the registration in person. Most of these areas were close to the border.

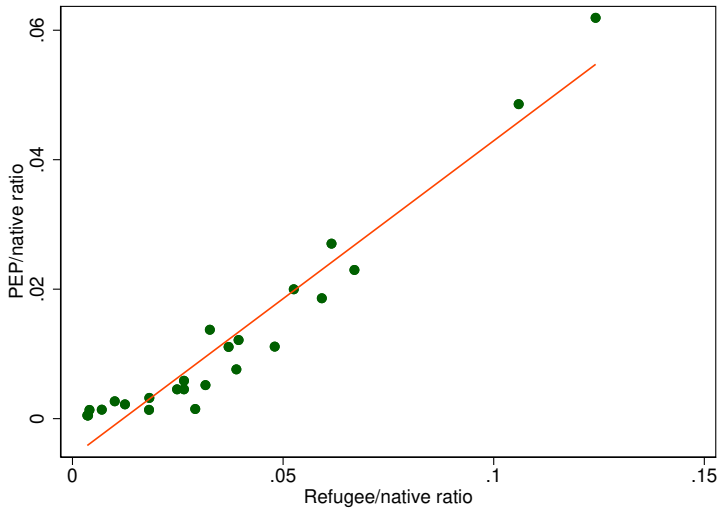
Decree 1288 of July 25th 2018, together with resolutions 6370 and 2033 of August 1st and 2nd, established that Venezuelans registered in the RAMV were also eligible for PEPs. This brought about the third wave of PEPs. Those registered in the RAMV were invited to apply between August 2nd and December 21st of 2018.

[17] All the laws, decrees, and resolutions can be found here

[18] Los pasaportes más caros y baratos de América Latina (y cuánto duran), BBC Mundo

PEPs were first of their kind in providing formal labor market access to migrants easily and at low cost However, anecdotally, they encountered a series of difficulties. The first, was that many employers initially did not understand that PEPs operated as work permits. There was a big push by the government to educate the public about their functionality following the first wave of PEPs Secondly, there were some technological challenges in their implementation. For example, the ID number provided by PEPs was shorter than the one used for Colombian national IDs. This meant that many systems could not process the migrant's new IDs. Finally, and most seriously, there began to be a lot of fraud. Since PEPs were simply printed from the migration office's website, some people began to replace the name of the recipient These considerations led the government to replace the PEP program with the PPT program. As part of this new program, IDs were issued using a more sophisticated set of biometrics.

Figure B.1: Spatial correlation between work permit treatment and refugee treatment intensity



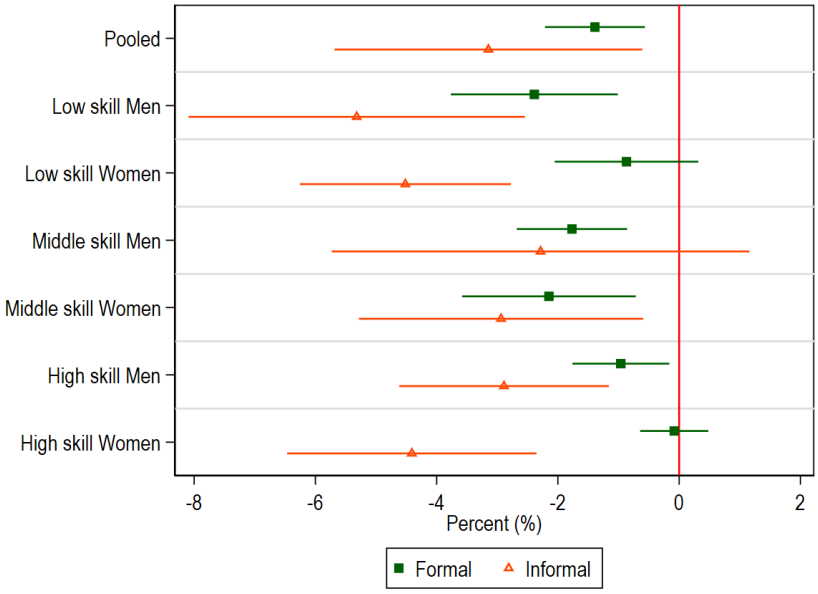
C Additional Results

Table C.3: 2SLS estimates reported in main text

Salaried Employment			Non-Salaried Employment			Employment		
All	Informal	Formal	All	Informal	Formal	All	Informal	Formal
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Pooled								
-0.417***	-0.267***	-0.144***	0.179	0.023	0.006	-0.142	-0.311*	-0.071
(0.111)	(0.110)	(0.044)	(0.173)	(0.169)	(0.073)	(0.275)	(0.183)	(0.112)
Panel B: Low-skill Men								
-0.729***	-0.633***	-0.177***	0.927***	0.962***	0.015	0.095	0.131	-0.297***
(0.129)	(0.169)	(0.052)	(0.355)	(0.343)	(0.028)	(0.328)	(0.309)	(0.076)
Panel C: Low-skill Women								
-0.203***	-0.209***	-0.017	0.275	0.113	0.061*	-0.158	-0.144	0.030
(0.044)	(0.041)	(0.012)	(0.497)	(0.454)	(0.037)	(0.488)	(0.492)	(0.042)
Panel D: Middle-skill Men								
-0.709***	-0.289	-0.335***	0.083	0.037	0.023	-0.634***	-0.437**	-0.212*
(0.245)	(0.223)	(0.088)	(0.157)	(0.156)	(0.095)	(0.263)	(0.198)	(0.110)
Panel E: Middle-skill Women								
-0.576***	-0.251***	-0.188***	0.234	0.220	0.051	-0.421***	-0.229**	-0.077
(0.068)	(0.102)	(0.064)	(0.200)	(0.160)	(0.043)	(0.104)	(0.109)	(0.087)
Panel F: High-skill Men								
-0.393***	-0.186***	-0.219***	-0.066	-0.377***	-0.074	-0.510***	-0.616***	-0.047
(0.129)	(0.057)	(0.093)	(0.125)	(0.093)	(0.107)	(0.158)	(0.106)	(0.107)
Panel G: High-skill Women								
-0.410***	-0.285***	-0.017	-0.069	-0.055	-0.048	0.182	-0.027	0.169
(0.089)	(0.068)	(0.058)	(0.173)	(0.160)	(0.065)	(0.316)	(0.142)	(0.261)

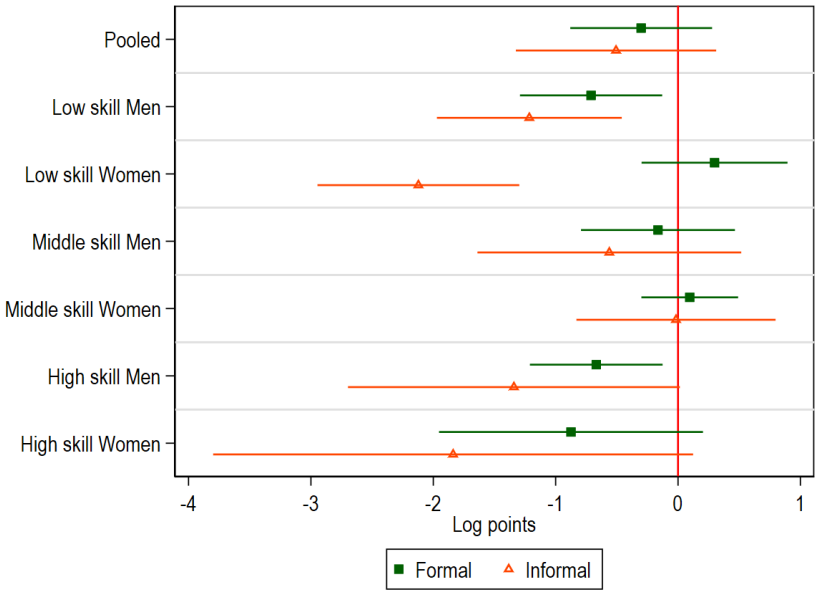
Note: Sample is retricted to those between ages of 15–65 in years 2007–2015. The estimates come from the Synthetic IV methodology described in text. Standard errors are clustered at the region level.

Figure C.2: LATE estimates in percentages



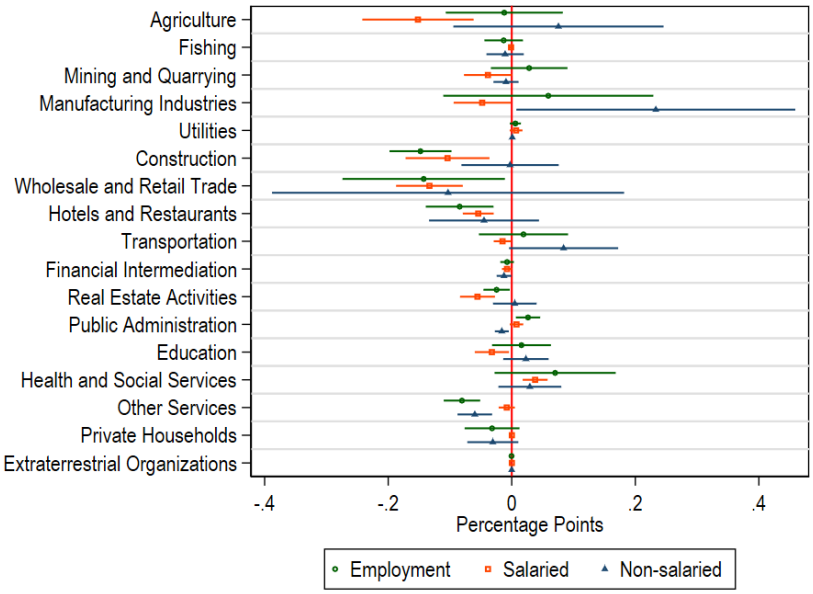
The estimates come from the Synthetic IV methodology described in text. The percentage effects on the informal and formal salaried employment rates across all skill levels are displayed. Standard errors are clustered at the region level.

Figure C.3: LATE estimates on wages



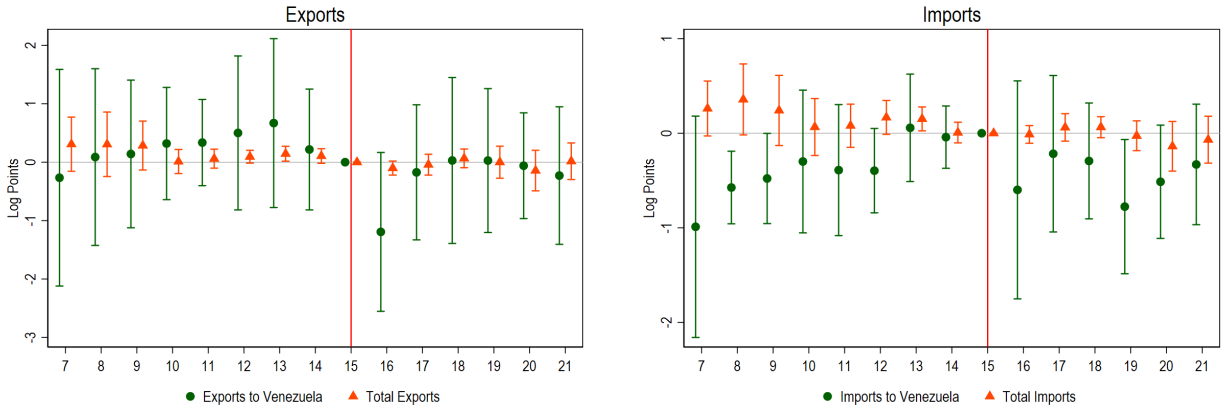
The estimates come from the Synthetic IV methodology described in text. The outcome is the natural logarithm of hourly wages of that skill group within salaried jobs. Standard errors are clustered at the region level. 95% confidence intervals are plotted.

Figure C.4: Heterogeneity by Industry



The estimates come from the Synthetic IV methodology described in text. Standard errors are clustered at the region level. Industry definitions follow ISIC standards.

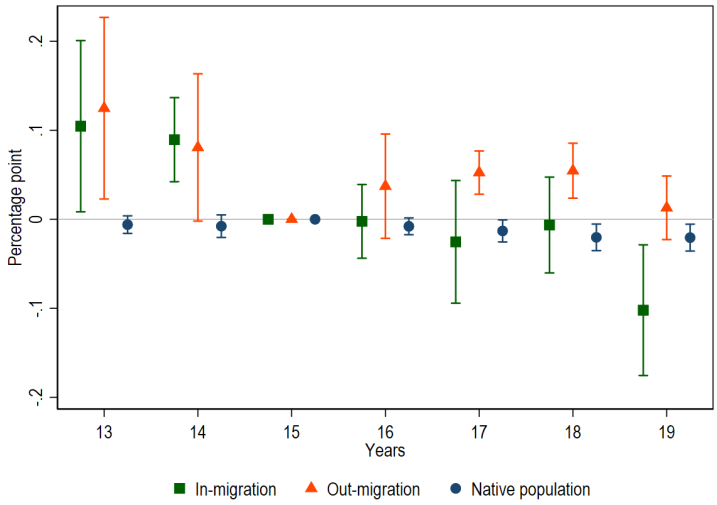
Figure C.5: The evolution of trade flows in the region. Correlation between distance and trade



(a) Exports

(b) Imports

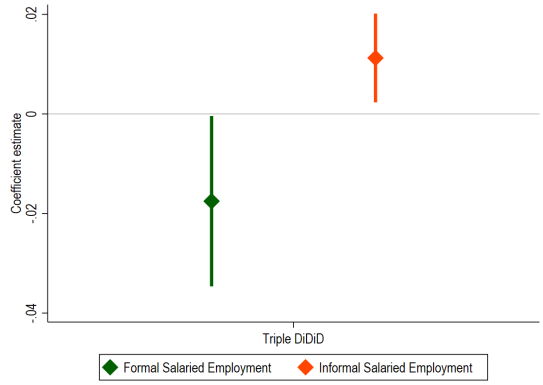
Figure C.6: Effect of distance on Internal migration



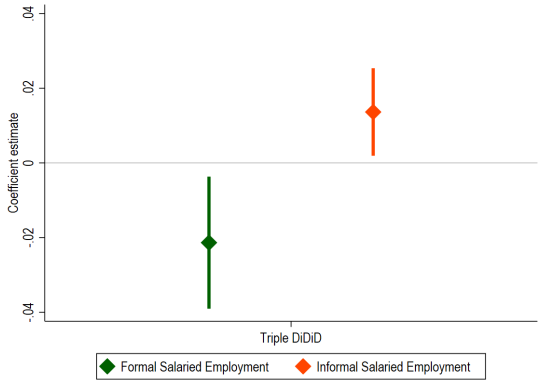
D Triple DiD Robustness Checks

This section presents the robustness checks for the triple DiD estimates on the effects of work permits. Figure D.7 shows the robustness to using different definitions of the regional treatment intensity and expanding the time-period of focus.

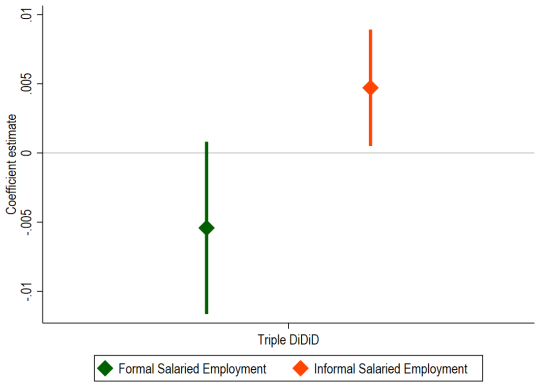
Figure D.7: Event-study estimates of education-sex exposure to work permits



(a) Main exposure indicator with time interval 2016q1-2019q4



(b) Alternative regional exposure



(c) Alternative regional exposure

Notes: Robustness check for Figure 8. Panel A expands the period from 2017q1-2019q4 to 2016q1-2019q4. Panel B changes the regional exposure definition from the 3 most exposed regions to the four most exposed regions. Panel C uses the continuous distance exposure definition. 95% confidence intervals are plotted.

We also estimate an event-study design to provide nonparametric estimates of our main design. While this allows for a visual check of our underlying assumptions, we lose some statistical power in the process due to having to estimate more coefficients.

We first compare the employment outcomes of high-skill native men to the remaining education-sex cells (omitting middle-skill men), before and after the provision of work permits in the third quarter of 2017, in regions close to and far away from the border (Equation 16).

We repeat this exercise for middle-skill native men (omitting high-skill men) before and after the fourth quarter of 2018 (Equation 17). This comparison can be written in an event-study model, separately for each treated cell, as follows:

$$Y_{brt} = \sum_{j \neq 2017q3} \beta_j \mathbb{I}_{b=HS-M} Z_r \mathbb{I}_{t=j} + f_{br} + f_{bt} + f_{rt} + \epsilon_{brt} \tag{16}$$

$$Y_{brt} = \sum_{j \neq 2018q4} \gamma_j \mathbb{I}_{b=MS-M} Z_r \mathbb{I}_{t=j} + f_{br} + f_{bt} + f_{rt} + \epsilon_{brt} \tag{17}$$

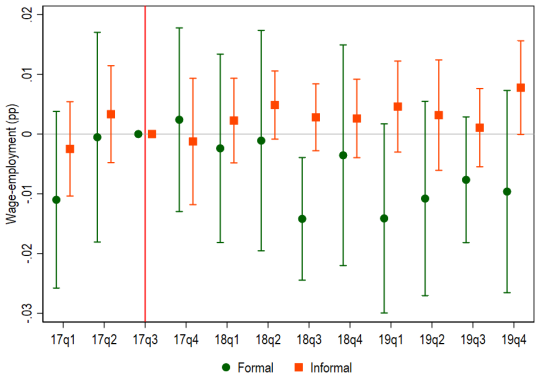
where f_{br} are bin*region fixed effects that capture anything about a region that might lead to differential outcomes by skill-cell, f_{bt} are bin*time fixed effects that capture any secular trends in labor market outcomes by skill-cell, and f_{rt} are region*time fixed effects that help us partial out regional trends. Crucially, under the assumption that the immigration shock itself does not impact the high-skill or middle-skill native men differently compared to other education-sex cells, these region*time fixed effects partial out the effect of immigration from the effect of work permits.

The identifying assumption is that, absent the provision of work permits, the gap between employment in regions close to and far away from the border, between treated men and others, would have evolved similarly over time. This is a statement, in differences, of the familiar parallel trends assumption. Visually, we would expect all the β_j and γ_j coefficients in equations 16 and 17 to be centered around zero absent treatment.

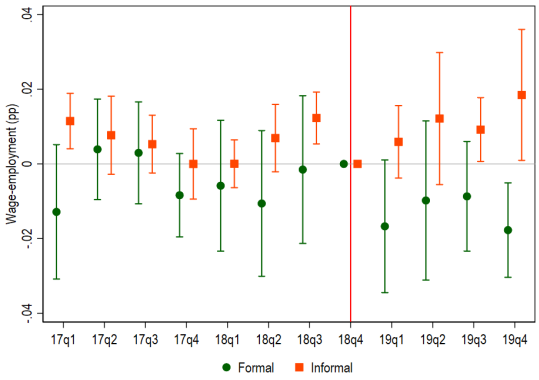
Additionally, the differential timing of the work permit treatment provides a sanity check for our identification strategy. We expect high-skill male natives to be impacted five quarters earlier than middle-skill male natives. Arguably, the likelihood that a confounder impacts high-skill and middle-skill men in the exposed regions five quarters apart is significantly low.

Figure D.8 plots the β_j and γ_j estimates in panels A and B respectively. Panel A shows that work permits reduce the formal salaried employment of high-skill men. They weakly increase, albeit insignificantly, their informal salaried employment. These effects are apparent by the third quarter of 2018, after the first two waves of PEPs. Panel B shows the same results for middle skill men. They too observe decreases in formal salaried employment and increases in informal salaried employment. Both results are statistically significant by the fourth quarter of 2019.

Figure D.8: Event-study estimates of education-sex exposure to work permits



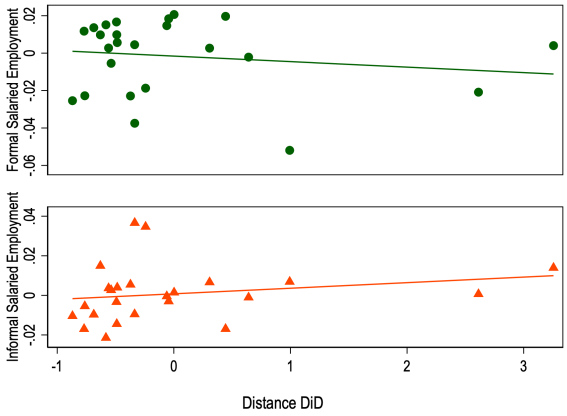
(a) High-skill men vs Control



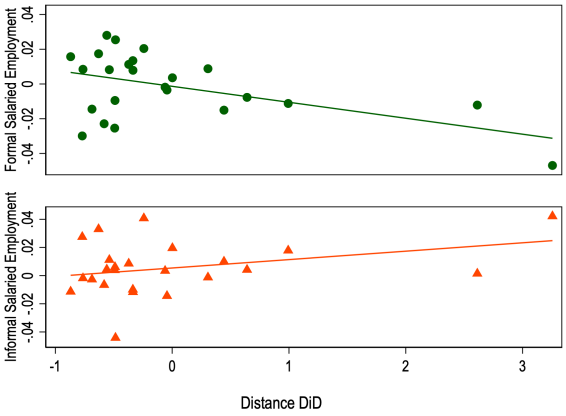
(b) Middle-skill men vs Control

The estimates come from triple DiD event-study designs: Panel A: $Y_{brt} = \sum_{j \neq 2017q3} \beta_j \mathbb{I}_{b=HS-M} Z_r \mathbb{I}_{t=j} + f_{br} + f_{bt} + f_{rt} + \epsilon_{brt}$ and Panel B: $Y_{brt} = \sum_{j \neq 2018q4} \gamma_j \mathbb{I}_{b=MS-M} Z_r \mathbb{I}_{t=j} + f_{br} + f_{bt} + f_{rt} + \epsilon_{brt}$. Control cells include low-skill men and all skill types of women. Standard errors are clustered at region-bin level.

Figure D.9: Skill-sex DiD within regions



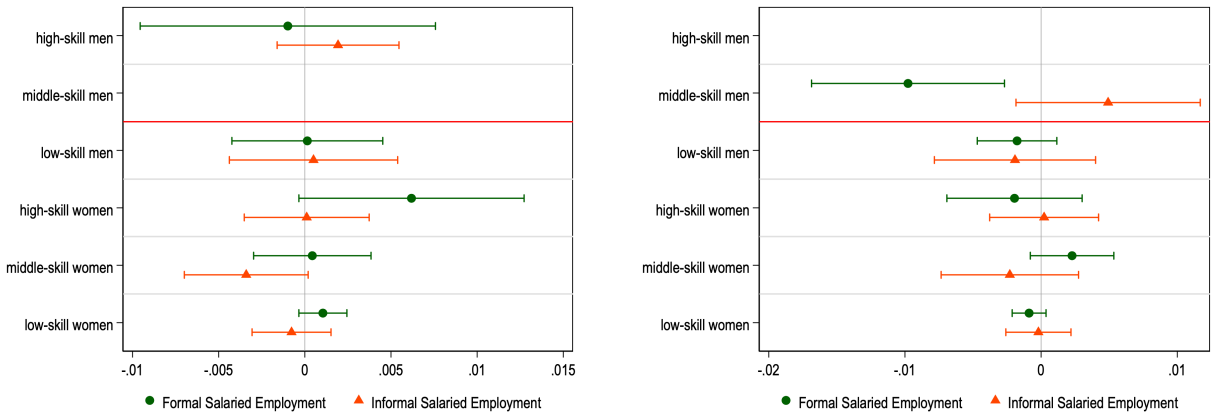
(a) High Skill Men vs rest



(b) Middle Skill Men vs rest

Notes: Each dot is a coefficient from a department-specific DiD estimation comparing employment outcomes across skill cells and time. To arrive at the triple DiD estimate we compare these estimates from regions closer to the border (left hand side of the x axis) to the estimates from regions further from the border (right hand side of the x axis).

Figure D.10: Spatial DiD within skill-sex groups



(a) High Skill Men vs rest

(b) Middle Skill Men vs Rest

Notes: Each row plots two coefficients from a skill-specific spatial DiD estimation comparing employment outcomes across regions close to and away from the border. Under the null that work permits had no impact, the estimates across rows would be the similar, i.e., lie on a vertical line. The triple DiD can be achieved by comparing the estimate of high-skill and middle-skill men against the average estimates of other skill-sex groups. Here, we also see that our triple DiD estimates are not driven by an outlier DiD estimate within the building blocks, for example, in one of the control groups.

E Model Estimation

This section discusses the estimation of the full model with firm heterogeneity. It closely follows Gulek (2023). We use a minimum distance estimator. Firm heterogeneity is introduced to obtain additional moments for identification. Section E.1 sets up the full model, while Section E.2 describes the estimation method, identification, and the model's fit.

E.1 Introducing Firm heterogeneity in productivity

Building on the representative firm framework of Section 3 we allow for firms to have different productivities denoted by $\theta \in \{\theta_1, \ldots, \theta_K\}$, which enters firms' production function in a Hicks-neutral way:

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

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

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

The first-order conditions determine the labor demand functions of each firm of type θ:

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

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

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

$$\begin{aligned} \overline{\epsilon_{L_i,w_i}} &:= \sum_{j=1}^{K} \epsilon_{L_i,w_i}(\theta_j) n_j \\ \overline{\epsilon_{L_f,w_i}} &:= \sum_{j=1}^{K} \epsilon_{L_f,w_i}(\theta_j) m_j \end{aligned}$$

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

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

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

We partition the vector of parameters into two groups based on whether they are calibrated or estimated. $\alpha = 0.49$ is calibrated based on the share of labor in production in Colombia (acquired from Penn World Tables), informal wage w_i and formal wage w_f for the low-skilled are estimated using the labor force surveys, the labor tax rate is set to its statutory value $\tau_w = 0.39$. The value of τ_w corresponds to the effective tax rate for minimum wage earners.

E.2 Estimation Method

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

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

$$\hat{\Phi} = \arg\min_{\Phi} Q(\Phi;\Psi) = \{g_N(\Phi;\Psi)' W_N g_N(\Phi;\Psi)\} \tag{18}$$

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

Moments and Identification

We use nine moments from the data and our SIV estimates to form the vector $\hat{m}_N$. GEIH asks respondents how many people work in their establishment, and group results in 7 categories: between 2–3, 4–5, 6–10, 11–19, 20–30, 31–50, and 51–100 workers. We follow this structure of the GEIH. The moments we choose are (i) the size of firms in different groups (calculated using GEIH), (ii) the informality rate of firms in different groups (calculated using GEIH), (iii) the ratio of informal and formal labor demand elasticities (estimated in the empirical section).

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

Estimates and Model Fit

Table E.4 shows the values of all parameters. The most critical estimate is that the CES elasticity parameter ρ is 0.91, which implies an elasticity of substitution between informal and formal labor of 11. To the best of our knowledge, this is one of the first papers to estimate this elasticity.[19] This relatively high elasticity is consistent with the Colombian context,

[19] There are two other papers that we are aware of that estimate this elasticity. Schramm (2014) studies the equilibrium effects of taxation on sectoral choice, work hours and wages in Mexico. She finds this elasticity to be around 1.8. Gulek (2023) studies the effect of informal Syrian refugees on Turkish labor markets and estimates this elasticity to be around 10. Informal and formal workers working in different sectors and firms in Mexico as opposed to working in the same firms as in Turkey and Colombia could explain this discrepancy.

Table E.4: Parameter Values

Parameter	Description	Source	Value
τ_w	Payroll tax	Statutory values	0.326
w_i	Informal wages	Calibrated	3.17
w_f	Formal wages for the low-skilled	Calibrated	5.949
α	Cobb-Douglass coefficient	Calibrated	0.49
γ	Intensive mg. cost of informal labor	Estimated	1.11
η	Informal share parameter	Estimated	0.86
ρ	CES elasticity parameter	Estimated	0.91
θ_1	Productivity of firms between 2–3 workers	Estimated	65
θ_2	Productivity of firms between 4–5 workers	Estimated	184
θ_3	Productivity of firms between 6–10 workers	Estimated	254
θ_4	Productivity of firms between 11–19 workers	Estimated	307
θ_5	Productivity of firms between 20–30 workers	Estimated	355
θ_6	Productivity of firms between 31–50 workers	Estimated	413
θ_7	Productivity of firms between 51–100 workers	Estimated	492
$\sigma_{i,f}$	Elasticity of substitution between informal and formal workers	Implied	11
ϵ_{L_i,w_i}	Elasticity of informal labor demand w.r.t. informal wages	Implied	-0.65
ϵ_{L_f,w_i}	Elasticity of formal labor demand w.r.t. informal wages	Implied	1.41

Note: Note: Formal and informal hourly wage (for salaried work) estimates are expressed in Pesos using the 2007–2015 sample of the GEIH.

where informal employment is often in the same sectors and even in the same firms as formal employment. It also supports the assumption of perfect substitutability between informal and formal workers in the recent structural literature on the informal sector (Ulyssea, 2018, 2020), and is quantitatively similar to the other estimate in the literature coming from Turkey (Gulek, 2023).

Appendix Table E.5 shows how the model performs compared to all of the targeted moments in the data. The model matches most of the moments of the data quite well. In general, it overestimates the informality of smaller firms and underestimates it for the larger firms.

Table E.5: Model Fit

Size of firm	Data	Model	
2–3 workers	GEIH	2.5	2.5
4–5 workers	GEIH	4.5	5.0
6–10 workers	GEIH	8.0	8.6
11–19 workers	GEIH	15.0	14.9
20–30 workers	GEIH	25.0	23.6
31–50 workers	GEIH	40.5	37.7
51-100 workers	GEIH	75.5	62.5
Share of informality			
2–3 workers	GEIH	0.87	1.00
4–5 workers	GEIH	0.74	0.94
6–10 workers	GEIH	0.57	0.65
11–19 workers	GEIH	0.39	0.41
20–30 workers	GEIH	0.28	0.27
31–50 workers	GEIH	0.18	0.18
51-100 workers	GEIH	0.13	0.11
Ratio of demand elasticities	SIV estimates	-0.44	-0.46

F Model explaining knowledge spillovers

The goal of this section is to rationalize the results from the quadruple difference in differences design in Section 5. There are R regions and J industries, each region-industry pair is populated with a representative firm, which uses low-skill and high-skill labor in production. For simplicity, we abstract away from informal and formal labor differences. Introducing a nested CES of informal and formal labor within each group would not change our predictions. Each firm's production function is given by

$$\begin{aligned} y_{rjt} &= f(H_{rjt}t, L_{rjt}, A_{H,rjt}, A_{L,rjt}) \\ &= A_{rjt}(A_{H,rjt}H_{rjt}^{\rho} + A_{L,rjt}L_{rjt}^{\rho})^{(1/\rho)} \end{aligned}$$

where A_{rjt} is a Hicks-neutral productivity measure for each firm, $A_{H,rjt}$ and $A_{L,rjt}$ are skill-enhancing productivity parameters for high and low-skill labor, respectively. Labor markets are competitive, workers are paid their marginal productivities. After some algebra, one can show:

$$ln(\frac{w_{H,rjt}}{w_{L,rjt}}) = ln(\frac{A_{h,rjt}}{A_{l,rjt}}) + (\rho - 1)ln(\frac{H_{rjt}}{L_{rjt}}) \tag{19}$$

This equation shows that the relatve wages of high-skill and low-skill workers in an industry-region cell depends on (1) the relative magnitudes of their productivitities and (2) their relative supplies.

In practice, both the relative technology measures and the relative skill intensities can evolve over time. For example, certain industries may become more productive over time, immigration can alter the skill ratios at the region level, etc. The empirical design controls for such changes by saturating the regression with three-way fixed effects.

Work permits can impact relative wages in two ways. First, absent work permits, high-skill refugees cannot work in certain industries due because of their high formality. Work permits enable refugees to alter their industry composition and work in more skill and formal intensive industries. This will change the relative supply of high and low-skill workers in a region-industry cell: the ratio of high to low skill labor in more formal intensive industries and in regions closer to the border will increase. Second, we acknowledge the possibility that the high-skill workers may not be purely a labor supply shock. Some refugees bring valuable human capital: specific know-how, ideas, etc. This decrease in mismatch can be rouphly captured in this model by a change in the relative productivities of high and low skill labor. In particular, we let the skill ratios and technologies in region-industry cells to evolve:

$$
\begin{aligned}
ln(\frac{H_{rjt}}{L_{rjt}}) &= \delta_{rt} + \delta_{jt} + \delta_{rj} + D_{rjt} + \epsilon_{rjt}, \\
ln(\frac{A_{H,rjt}}{A_{L,rjt}}) &= g_{rt} + g_{jt} + g_{rj} + \theta_{rjt} + \eta_{rjt},
\end{aligned} \tag{20}
$$

where the fixed effects δ and g allow the skill ratios and productivities to differ at the region-year, industry-year, region-industry level; D_{rjt} is the change in the effective skill ratios due to work permits, θ_{rjt} allows for the possibility that high-skill refugees impact the technology parameters in the exposed industry-region cells. To get our main quadruple DiD design, plug in these terms into the equation to get:

$$
ln(\frac{w_{H,rjt}}{w_{L,rjt}}) = \underbrace{f_{rt} + f_{jt} + f_{rj}}_{\text{F: Fixed effects}} + \underbrace{(\rho - 1)D_{rjt} + \theta_{rjt}}_{\text{E: exposure terms}} + \epsilon_{rjt} \tag{21}
$$

In practice, if we had only two skill groups, we could run this regression at region-industry-year level. However, we show that work permits impact high-school graduate and college graduate refugees at different times, which we exploit for additional statistical power. To obtain a regression equation at skill-region-industry-time level, we write:

$$
\begin{aligned}
ln(\frac{w_{H,rjt}}{w_{L,rjt}} &= \underbrace{f_{rt} + f_{jt} + f_{rj}}_{\text{F}} + \underbrace{(\rho - 1)D_{rjt} + \theta_{rjt}}_{\text{E}} + \epsilon_{rjt} \\
&\Rightarrow ln(w_{H,rjt}) = ln(w_{L,rjt}) + F + E + \epsilon_{rjt} \\
&\Rightarrow ln(w_{H,rjt}) = f_{rjt} + F + E + \epsilon_{rjt} \\
&\Rightarrow ln(w_{rjt}) = f_{rjt} + \mathbb{1}(skill = h)(F + E) + \epsilon_{rjt} \\
&\Rightarrow ln(w_{srjt}) = f_{rjt} + f_{srt} + f_{sjt} + f_{srj} + (\rho - 1)D_{srjt} + \theta_{srjt} + \epsilon_{srjt}
\end{aligned} \tag{22}
$$

where we let f_{rjt} to capture the wages of the baseline skill-level workers (i.e., low-skill workers), and define $D_{srjt} = D_{rjt}$ if $s = h$, 0 otherwise.

If work permits do not create any change in the technology through knowledge spillovers, then $\theta_{srjt} = 0$. In this case, we should see a decrease in relative wages of middle to high skill male natives, in more skill/formal intensive industries, in regions closer to the border, after the granting of work permits. The fact that we observe a positive slope in Figure 9 suggests that $\theta_{srjt} > 0$, meaning that the high-skill refugees make the high-skill workers in the exposed region-industries more productive compared to low-skill workers. This is one way to interpret what we refer to as the knowledge spillover from reduction of skill mismatch

in the main text.

G Model Appendix

G.1 Empirical concavity of non-salaried wages

Table G.6: The concavity of home production for different groups of natives

	Men	Women	Low-skill	Middle-skill	High-skill
hours	0.0437***	0.0671***	0.0530***	0.0525***	0.0659***
	-0.000242	-0.000213	-0.000195	-0.000267	-0.000395
$hours^2$	-0.000282***	-0.000507***	-0.000344***	-0.000346***	-0.000543***
	-0.00000203	-0.0000024	-0.00000184	-0.00000248	-0.00000425
N	761990	636745	710154	351996	336585

Note: The outcome is the natural logarithm of monthly income in all columns. Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

H Synthetic Instrumental Variables

In this section, we explain briefly the idea behind the Synthetic IV methodology and how it works. we refer the reader to Gulek and Vives-i Bastida (2023) for a full treatment.

The synthetic IV (SIV) is a non-parametric method that combines instrumental variable strategy with synthetic controls. SIV partials out the unmeasured confounding in IV-DiD settings such as exposure or shift-share designs, and under standard conditions, is asymptotically normal when the standard two-stage least squares is not. We employ SIV in two steps.

Step 1: For each region i, we find synthetic control weights w_{ij} by solving the standard synthetic control program for past outcomes Y_{jt}. We use these weights to generate the synthetic outcome Y^{SC}, treatment R^{SC}, and instrument Z^{SC}

$$\begin{aligned}
\hat{Y}_{it}^{SC} &= \sum_{j \neq i} \hat{w}_{ij}^{SC} Y_{jt}, \\
\hat{R}_{it}^{SC} &= \sum_{j \neq i} \hat{w}_{ij}^{SC} R_{jt}, \\
\hat{Z}_{it}^{SC} &= \sum_{j \neq i} \hat{w}_{ij}^{SC} Z_{jt},
\end{aligned}$$

Then, we compute the *debiased* values.

$$\begin{aligned}
\tilde{Y}_{it} &= Y_{it} - \hat{Y}_{it}^{SC}, \\
\tilde{R}_{it} &= R_{it} - \hat{R}_{it}^{SC}, \\
\tilde{Z}_{it} &= Z_{it} - \hat{Z}_{it}^{SC}.
\end{aligned}$$

Step 2: We estimate the first-stage and the reduced-form using the debiased data.

The intuition behind SIV is simple. Step 1 creates a synthetic "control" for all the regions in the data that follow similar trajectories before the refugee shock begins. This addresses the "pre-trend" problem. However, it does not solve the problem that immigrants can choose their location based on contemporaneous economic shocks. This is still addressed by the instrument Z. Put differently, SIV addresses both the unobserved confounding problem via SC and the endogeneity problem via IV by combining both.

H.1 SIV Quality Checks

In this section we describe the robustness checks we perform for the SIV methodology. Gulek and Vives-i Bastida (2023) suggest three sets of checks to validate the robustness of the SIV

methodology. We perform the robustness checks of the synthetic IV methodology suggested by Gulek and Vives-i Bastida (2023):

1. **Check your first stage**: "given that the debiasing procedure can lead to a weaker first stage, in cases with strong omitted variable bias if the synthetic IV estimator exhibits a weak first stage researchers may be worried using the synthetic estimator."

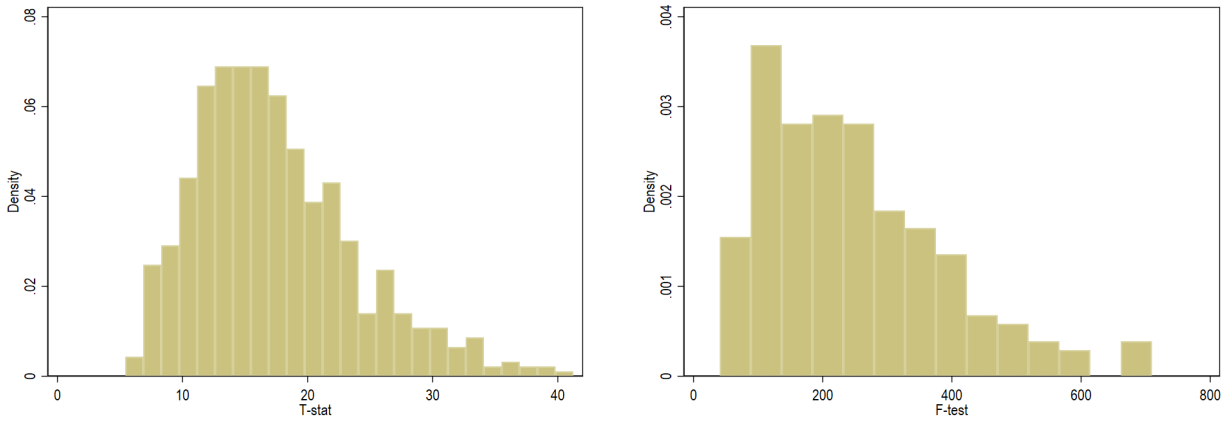
2. **Check your pre-treatment fit**: "if the debiased outcomes exhibits large deviations in the pre-treatment period or an event study design reveals pre-trends, it is likely that the synthetic estimator will be biased."

3. **Back test**: "given that the finite sample bias depends on the expected pre-treatment fit, back testing can reveal whether good pre-treatment fit was due to over-fitting (biasing the estimator) or not."

We perform all these robustness checks and report the results here.

First, we need to ensure that a strong first stage remains after debiasing. For each of the 81 outcomes reported in Figure 7, we estimate equation 9 using debiased data after applying both baseline SIV and SIV with demeaned outcomes as explained in the text. Figure H.11a shows the t-stats of the individual coefficients of the instrument interacted with the year dummies in 2017, 2018, and 2019, and Figure H.11b shows the joint significance of these three coefficients for each specification. Our debiasing procedure not only maintains the positive first-stage coefficients, i.e., (debiased) inverse distance predicts more (debiased) refugee treatment, but the coefficients remain significant, with the smallest t-stat being 5 out of 486 estimates, and the smallest joint F-statistic being 59 with the 5th percentile is 98.

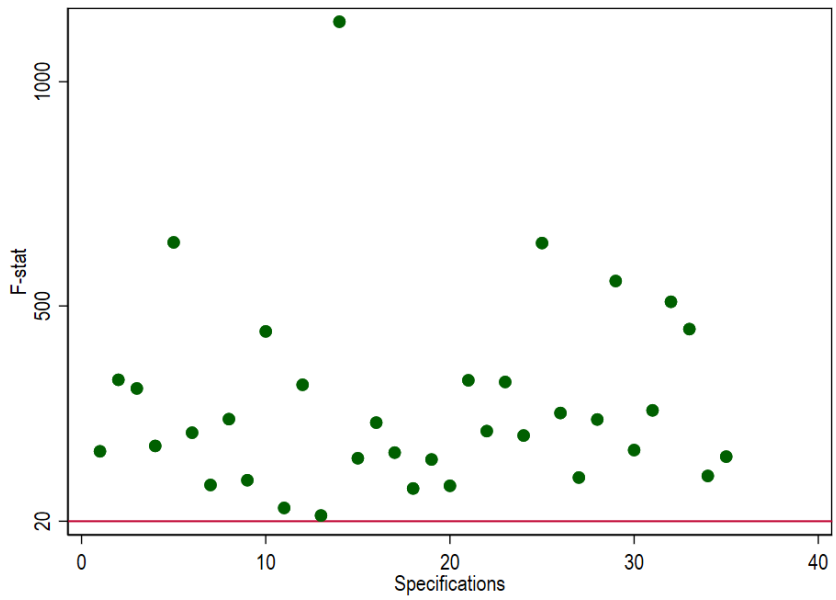
Figure H.11: First-stage statistics after debiasing



(a) Individual T-stats

(b) F-test

Figure H.12: F-statistic of the debiased instrument



F-statistic of the instrument is calculated by estimating the equation $\widetilde{R_{it}} = \beta \widetilde{Z_{it}} + f_i + f_t + \eta_{it}$, where $\widetilde{R_{it}}$ and $\widetilde{Z_{it}}$ are the debiased versions of the treatment and instrument, respectively. f_i and f_t are region and time fixed effects. Standard errors are clustered at the region level. F-statistic of the instrument $\widetilde{Z_{it}}$ from the 35 main specifications shown in Figure 7 are plotted.

Recall that we rely on one shift-share instrument Z_{it} for one treatment R_{it}. Consequently, in each regression, we obtain a separate first-stage coefficient. Figure H.12 shows the scatterplot of these first-stage estimates. Of the 35 specifications, the minimum F-statistic of the instrument in the post-period (2016–2019) is 32, with more than 75% of the specifications having an F-stat of more than 100. Overall, we maintain a strong first-stage in all of our main estimates.

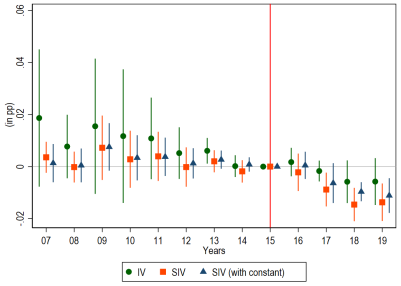
Second, we present the event-study figures using IV and SIV for all the outcomes reported in Figure 7. Figures H.14, H.15, and H.16 present the event study estimates on salaried, non-salaried, and overall employment rate of natives using IV and the two versions of SIV.

Third, we match on deviations from pre-treatment mean instead of levels in the synthetic control problem and repeat the main analyses. For almost all outcomes, this provides qualitatively similar results to baseline SIV.

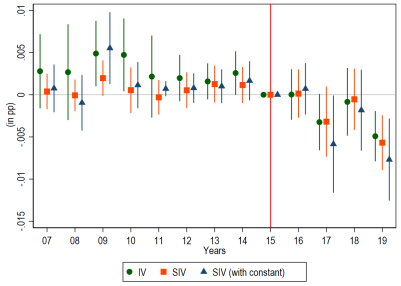
Fourth, we perform backtesting. Instead of using the 2007–2015 values of the outcome in the matching process, we use values from 2010–2015. Figure H.18 compares results reported in Figure 7 reported in the main text with the version using a different training set. Overall, results are highly comparable across the two specifications.

Figure H.13: Event-study estimates of Immigrants' impact on salaried, formal and informal employment of natives (IV vs SIV vs SIVdm)

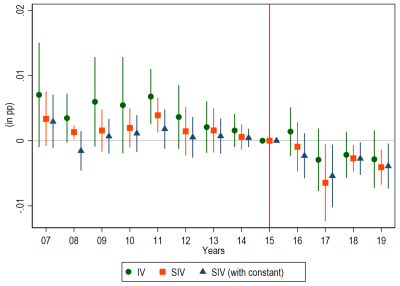
(a) Informal: low-skill men



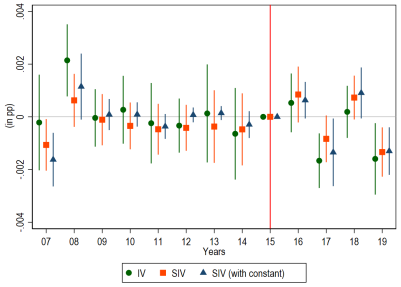
(b) Formal: low-skill men



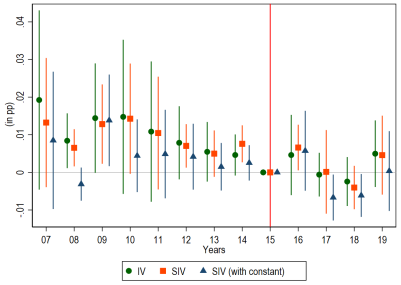
(c) Informal: low-skill women



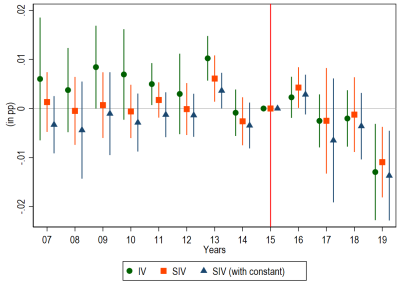
(d) Formal: low-skill women



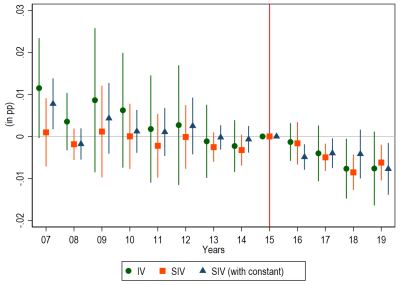
(e) Informal: middle-skill men



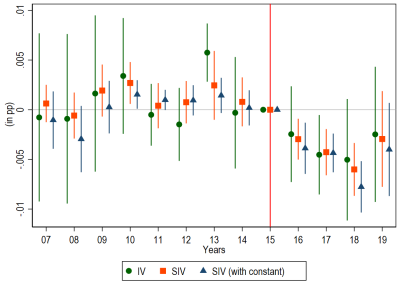
(f) Formal: middle-skill men



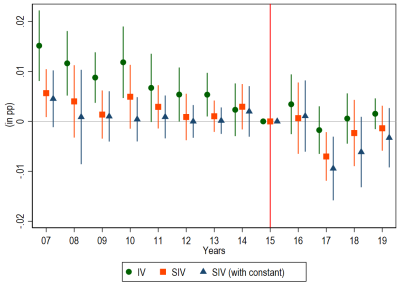
(g) Informal: middle-skill women



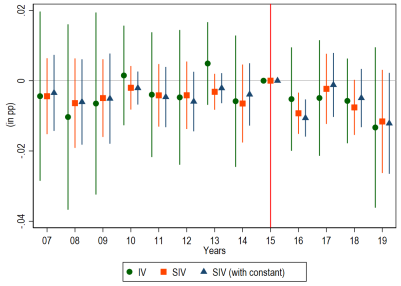
(h) Formal: middle-skill women



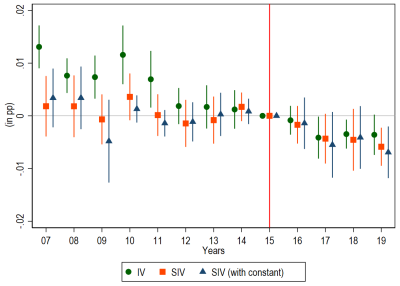
(i) Informal: high-skill men



(j) Formal: high-skill men



(k) Informal: high-skill women



(l) Formal: high-skill women

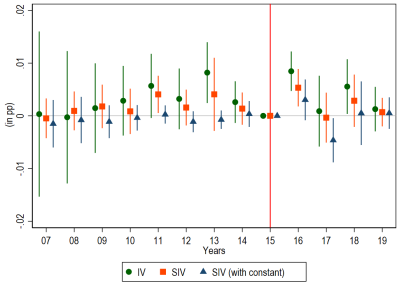
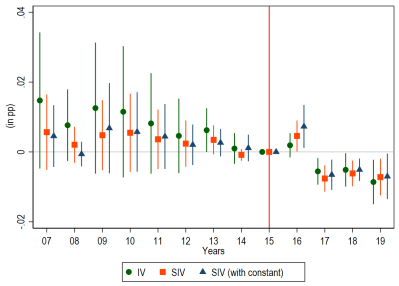
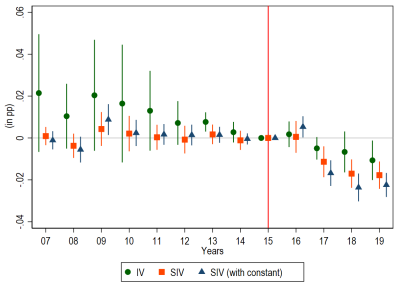


Figure H.14: Event-study estimates of Immigrants' impact on salaried employment of natives (IV vs SIV vs SIVdm)

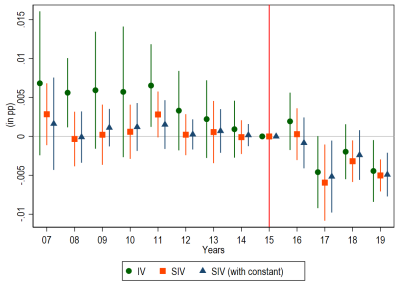
(a) Pooled



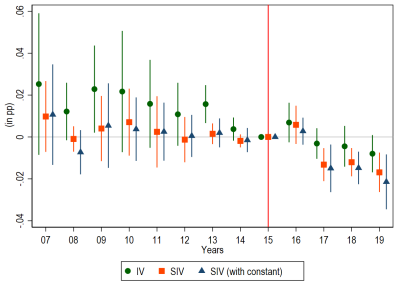
(b) Low-skill Men



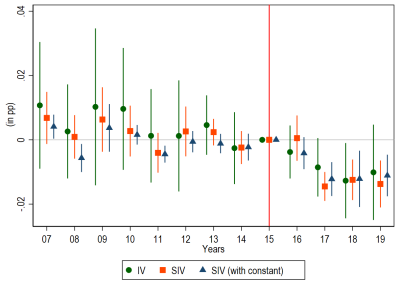
(c) Low-skill Women



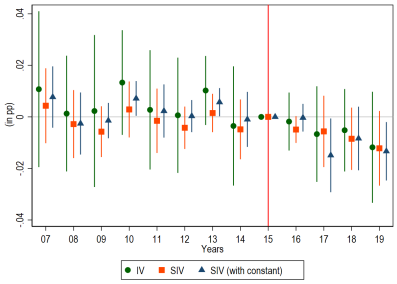
(d) Middle-skill Men



(e) Middle-skill Women



(f) High-skill Men



(g) High-skill Women

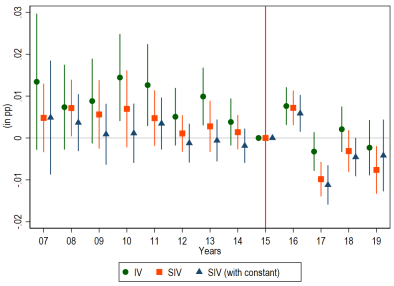
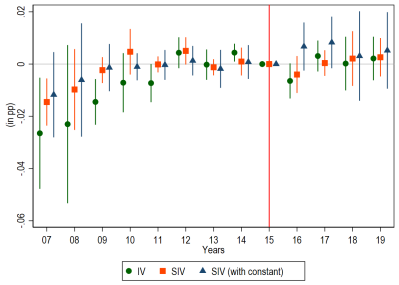
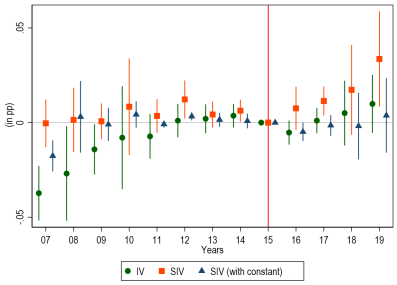


Figure H.15: Event-study estimates of Immigrants' impact on non-salaried employment of natives (IV vs SIV vs SIVdm)

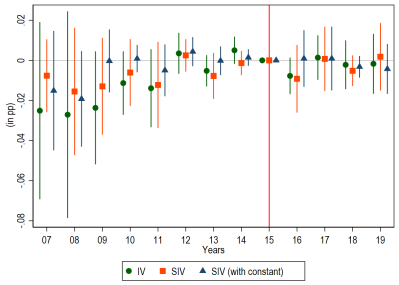
(a) Pooled



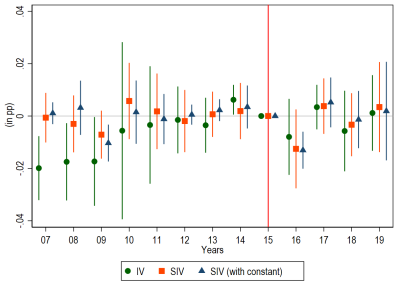
(b) Low-skill Men



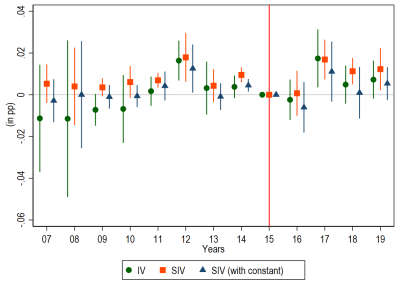
(c) Low-skill Women



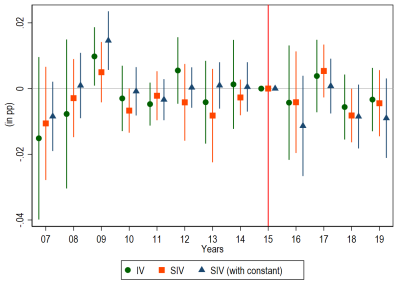
(d) Middle-skill Men



(e) Middle-skill Women



(f) High-skill Men



(g) High-skill Women

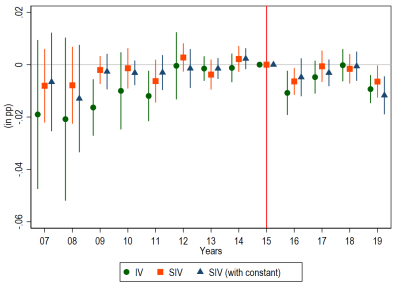
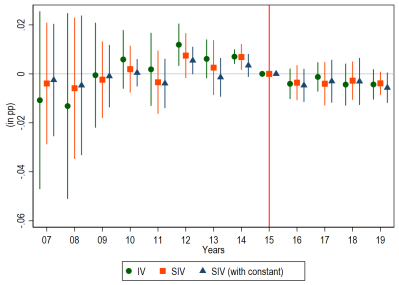
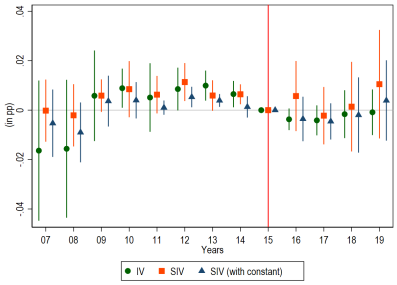


Figure H.16: Event-study estimates of Immigrants' impact on employment of natives (IV vs SIV vs SIVdm)

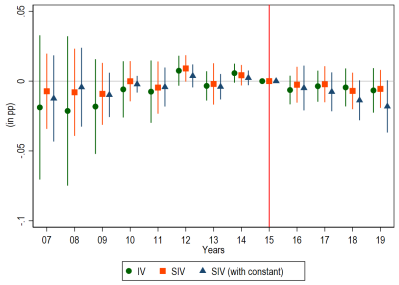
(a) Pooled



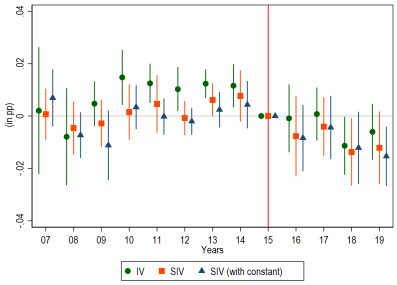
(b) Low-skill Men



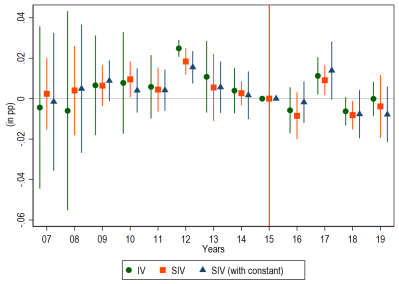
(c) Low-skill Women



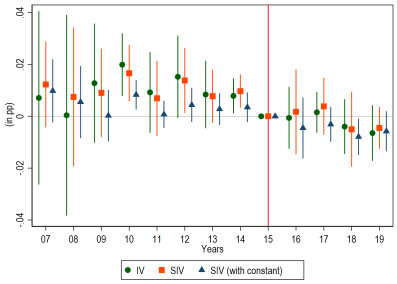
(d) Middle-skill Men



(e) Middle-skill Women



(f) High-skill Men



(g) High-skill Women

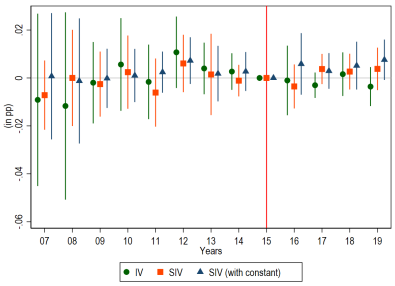
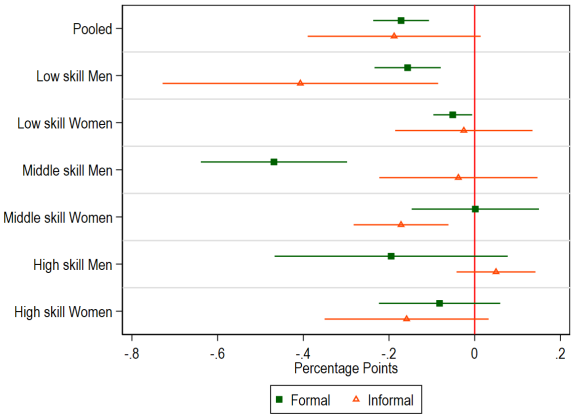
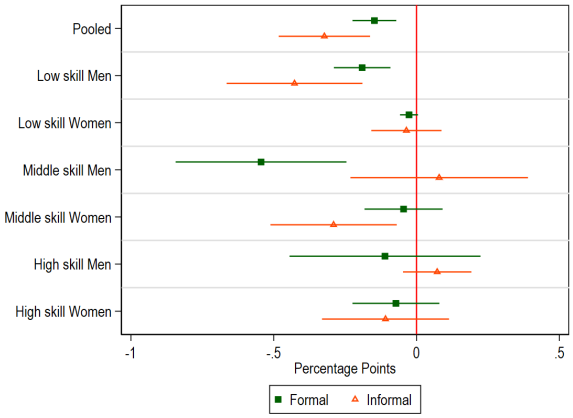


Figure H.17: Robustness of Figure 7a

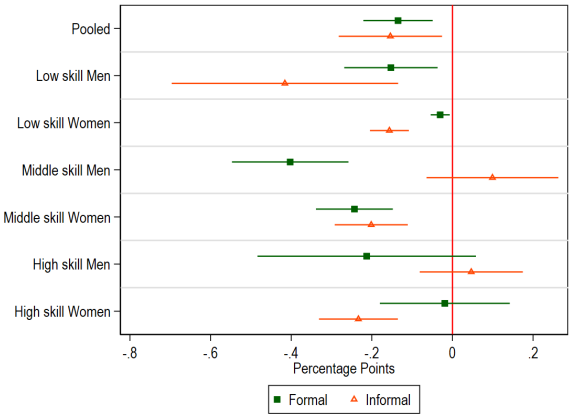
(a) Training: 2007–2015



(b) Training: 2010–2015



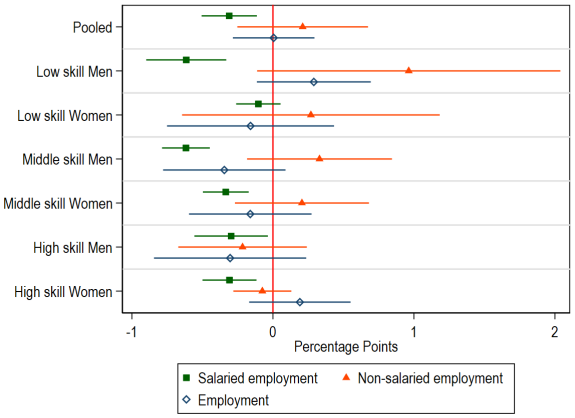
(c) Training: 2007–2012



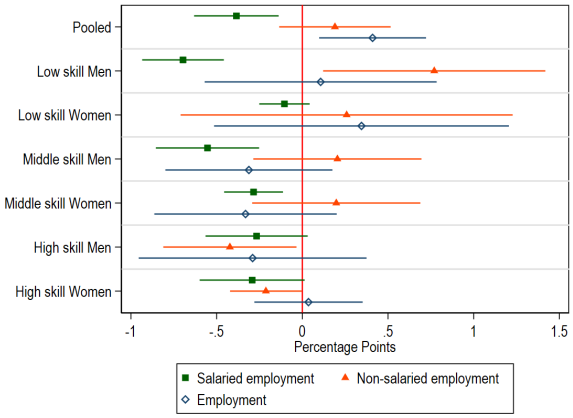
The estimates come from the Synthetic IV methodology described in text. Standard errors are clustered at the region level.

Figure H.18: Robustness of Figure 7b

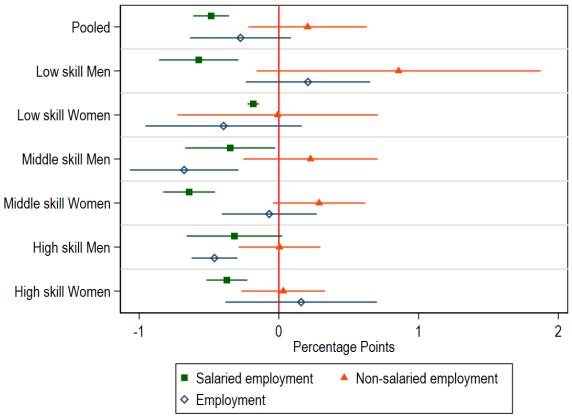
(a) Training: 2007–2015



(b) Training: 2010–2015



(c) Training: 2007–2012



The estimates come from the Synthetic IV methodology described in text. Standard errors are clustered at the region level.