Article

Impact of Venezuelan Migration on the Informal Workforce of Native Workers in Colombia

William Prieto Bustos [1,*]**, Cristian Darío Castillo Robayo** [2]**, Jacobo Campo Robledo** [1]
and Juliana Molina Dominguez [1]

[1] Faculty of Economic and Administrative Sciences, Catholic University of Colombia, Street 47 No. 13-54, Building C, Bogotá 110211, Colombia; jacampo@ucatolica.edu.co (J.C.R.); jmolina92@ucatolica.edu.co (J.M.D.)

[2] Faculty of Economic and Administrative Sciences, Los Llanos University, Street 37 No. 02-41, Villaviencio 1745110, Colombia; cristian.castillo.robayo@unillanos.edu.co

[*] Correspondence: woprieto@ucatolica.edu.co

Abstract: Colombia experienced a substantial increase in annual migration flow from Venezuela from 2013 to 2019, accumulating 1.7 million migrants by the end of 2019. According to migration data, 2016 was a breaking point in migration growth, with an increase of 137.57% compared with 2015 and at which time the influx of migrant workers began to be massive, rapid, and involuntary. In this regard, the research paper investigates, using a difference-in-differences model, the impacts on the labor market across different definitions of the informal workforce, testing the hypothesis that short-term labor migration increased (1) the number of employed individuals in companies with a workforce of fewer than five people, (2) the number of employed individuals not contributing to the social security system, and (3) the relative participation of the informal workforce in total employment from 2015 to 2018. The main results indicate an expansion in the labor market's informal segment, increasing the number of non-returned native workers in the informal workforce without significant increases in the participation of informality in total employment. The results remain robust across various samples in models adjusted for departmental-, municipal-, and individual-level data. Following the economic theory, the research findings seem to follow a transmission mechanism in which migrant workers reduce labor costs and increase production in informal markets, providing better conditions to increase informal jobs for native workers. Several national and international stakeholders implementing income-generation alternatives in the border departments focusing on migrant employment services could find the research findings helpful in at least two aspects: (1) fighting cultural stereotypes upon which basis native workers tend to see migration as a threat to their current job holdings requires evidence that shows migrant workers contribute to economic growth and employment; (2) promoting better public policies to take advantage of initial conditions that favor labor integration of migrant workers such as cultural and language similarities among natives and migrants works better when there is evidence of the migration's positive impacts.

Keywords: migration; labor markets; informal workforce labor markets

1. Introduction

The rapid and forced migration from Venezuela occurred massively from 2013 to 2019. According to data from the International Organization for Migration (IOM), more than 2.3 million people had left Venezuela for Colombia (26%), Chile (5.17%), and Argentina (2.48%) in Latin America. Recent studies on migration from Venezuela have shown an increasing impact on Colombia, Peru, and Ecuador's labor markets.

Within the migratory flow that affected Colombia, there were approximately 300,000 repatriated Colombian nationals, 468,428 Venezuelan nationals with regular migratory status, and 105,766 with irregular migratory status. Migration to Colombia of Venezuelan citizens and Colombian citizens who lived in Venezuela has increased significantly since 2015, with an annual increase of 126%, compared to a 43% increase in 2014.

Annual increases in migration reached 174% and 188% in 2017 and 2018, respectively, which initially increased the participation of migrants in border departments. According to the identified mobility pattern, in the first stage, migrants enter Colombia through regular points with the presence of Colombian Migration and through irregular points along the 2219 km of land border shared by both countries. In the second stage, the migratory flow, with more information about the destination and intentions to stay in Colombia, moves to urban centers with greater income-generating possibilities in Colombia. In contrast, those who decide not to stay in Colombia continue moving south towards the Ecuadorian border or north through the Darien Gap to enter Central America through Panama.

According to the diagnosis carried out by the World Bank, the migratory flow into the Colombian labor market consists of young migrants with educational levels relatively similar to those of unskilled native workers in the host regions. On the other hand, two stages are relevant to understand migration. In the first stage, the migratory flow reaches the capital cities of the border departments. In the second stage, it moves to urban centers with better employment opportunities and income generation. At the same time, information networks within the migratory flow affect movements in the second stage. In contrast, in the first stage, the flow is random, reaching border cities or cities near the Venezuelan border, such as Cúcuta, Arauca, Bucaramanga, and Riohacha.

The data source for constructing informal workforce indicators is the Integrated Household Survey (IHS). The IHS is a survey oriented toward labor supply and implemented by the National Administrative Department of Statistics (NADS) to monitor the labor market. The survey respondents are a sample of households representative of departments and city populations. The IHS surveys every household member regarding labor status during the previous week of the interview. Having an unbalanced panel of individuals instead of a longitudinal balanced panel of the same individuals becomes a limitation as long as the sample average of control variables could move along with the sample of individuals interviewed. However, descriptive analysis of the model variables shows a slight change in the control variables, proving grounds to consider that the sample design of the IHS survey fits the econometric model requirements.

The assembly of the informal workforce indicators' databases was constructed for three different units of observation: (1) departments, (2) capital cities, and (3) individuals. The database feeds the econometric specifications to estimate impact models on the informal workforce at the departmental and metropolitan levels of capital cities and a non-linear model of changes in the probability of being in the informal workforce at the individual level.

The findings suggest an expansion in the informal labor market segment in the treatment group's departments and cities, controlling for educational level and work experience. In the informal workforce measured by company size, migration increased the number of informal workers by 107,933 in the departmental model. In the city model, there was an increase of 35,139 informal workers. Both results are statistically significant. On the other hand, in the individual model, migratory flow increased by 2.5% in informal labor markets. The result is significant at the 1% level.

The results are similar to those using the definition of the informal workforce associated with employment that does not contribute to the social security system. In the departmental model, controlling for education and work experience, migratory flow during the first phase increased the number of informal workers by 361,966 compared to the control group's counterfactual. In the capital city model, the impact of migratory flow amounted to 118,696 individuals, remaining statistically significant at the 1% level. Finally, migratory flow without control variables (education levels and work experience) registered a positive impact of 2.5 percentage points on individuals' informality in the treatment group compared to the counterfactual. However, including control variables reduces the probability of informality and statistical significance. The reason for this result is associated with the positive correlation between education levels and work experience with contributions to the social security system. Thus, lower education levels and work experience

result in higher informality due to lower labor productivity, which reduces labor income and social security contributions.

Finally, the research presents results related to the informal workforce-to-employment ratio for the definitions of informality by company size and informality due to the absence of social security contributions. The informality-to-employment ratio, a measurement of the percentage participation of informality, increases in the departmental and municipal adjustments without reaching statistical significance. The same occurs in the individual adjustment, with positive percentage increases but without statistical significance.

Research findings suggest a positive response in informal jobs for native workers without a significant change in the number of native informal workers related to total employment. Following the economic theory, the research findings seem to follow a transmission mechanism in which migrant workers reduce labor costs and increase production in informal markets, providing better conditions to increase informal jobs for native workers. Hiring more qualified or equally qualified informal migrant workers at a lower wage increases the opportunities to hire native workers because they increase labor productivity over labor costs at a given level of demand for informal production. Such an improvement in informal production technology causes an expansion of the informal segment that provides a case for showing the effects on labor productivity as a primary driver of an increase in informal job opportunities for native workers. Moreover, total employment also increased due to the migration of Colombian nationals that lived in Venezuela, making the proportion of informal workers to total employment remain without significant change.

Unlike the recent literature on migration impacts in Latin America, which focuses on the second phase of migration to cities with more significant employment opportunities and income generation, this paper emphasizes the impact of migration on the informal workforce in labor markets in bordering departments, considering their economic development challenges. Therefore, this study's primary contribution lies in applying a methodology to measure the initial impact of migratory flow on the informal workforce in bordering departments, which initially face constraints in complementing the migrant flow.

There are limitations to research findings related to measured long-term impacts due to the econometric approach. First, the difference-in-differences (DID) model uses an unbalanced panel data method with fixed effects to control for characteristics that influence the migration effect but remain constant over time. In this regard, there is no explicit link to economic growth gains due to migration in the long term that might cause sustainable labor productivity changes that benefit native workers. In this regard, it only captures short-term effects that could vanish or prevail in the long term. Therefore, effects on consumption and investment caused by migration that might increase or decrease job opportunities for native workers are not in the research scope. However, the evidence of short-term effects could help inform public policy designs relevant to promoting migration's positive effects on economic welfare.

Seven sections structure the paper besides the introduction. The second section presents a literature review on rapid and massive migration's effects on labor markets. The third section describes the migration flow demographics. The fourth section explains the methodology and data sources used with a descriptive analysis of the variables of interest in the treatment and control groups. The fifth section describes the econometric model results, indicating the migration's effect on informal jobs in departments, capital cities, and individuals in the treatment group. The sixth section discusses the results, and the last section provides the conclusions.

2. Literature Review

The reviewed literature on rapid and massive migrations has pointed out evidence regarding the impact of migratory flows on the informal workforce and wages (Mora et al. 2022; Santamaria 2020; Peñaloza-Pacheco 2022; Pedrazzi and Peñaloza-Pacheco 2021; Caruso et al. 2019; Morales and Pierola 2020; Olivieri et al. 2022; Esen and Binatli 2017; Bağır 2018; Tumen 2016;

Fallah et al. 2019). According to the findings, the effects are positive on the informal workforce, inconclusive regarding wages, and negative on the employment of unskilled labor, particularly affecting vulnerable populations facing barriers to labor market access. At the same time, the long-term effects on productivity, consumption, fiscal income, and foreign trade have not been explored with the same frequency, as they involve a production function approach (Aydemir and Borjas 2007; Llull 2018; Borjas 2014; Card and Peri 2016; Ottaviano and Peri 2021) different from the impact measurement approach using difference-in-differences models (Altonji and Card 2019; Card 1990, 2001; Balkan and Tumen 2016; Ceritoglu et al. 2017; Mora et al. 2022; Morales and Pierola 2020; Olivieri et al. 2022; Peñaloza-Pacheco 2022; Pedrazzi and Peñaloza-Pacheco 2021; Santamaria 2020; Esen and Binatli 2017; Bağır 2018; Fallah et al. 2019; Caruso et al. 2019). Becker and Ferrara (2019) review the research agenda on rapid and massive migration's effects relevant to identifying the main challenges when estimating labor market effects.

Despite the above, the exploration of the impact on the informal workforce and wages, as well as the effect on the informal workforce and wage distribution, is relevant from a methodological perspective because the migratory flow in the first phase is random (Tumen 2015; Becker and Ferrara 2019). Even though there may be social, economic, and cultural reasons related to endogenous information that affects the decision-making process before migrating, the massive and rapid displacement reflects an unstructured and immediate decision-making process, different from a voluntary and planned migration decision (Borjas 2003; Aydemir and Borjas 2007; Llull 2018; Borjas 2014; Ottaviano and Peri 2021).

The previous condition creates an optimal environment for measuring the impact of migratory flows on the labor market in border departments, the cities closest to the borders, and on native informal workers residing in these territories because there is no correlation between the decision to migrate and the destination territory (Tumen 2015) in involuntary migrations. On the other hand, the second phase of the migratory flow is more endogenous due to the existence of information networks that guide the movements of migrants toward destinations with better employment opportunities and income generation. This has been modeled in recent literature using instrumental variable schemes that utilize the distance between Venezuelan states and Colombian departments (Pedrazzi and Peñaloza-Pacheco 2021), climate (Mora et al. 2022), and internet use (Santamaria 2020) as exogenous variables to control for endogeneity problems, which can cause biases in estimating the coefficients measuring the impact on the informal workforce, wages, and employment.

The migrant's two-phase approach relies on the assumption that moving out from the border cities is costly without high-quality information about potential income generation opportunities. Given that the migration flow was rapid and occurred in a short period, and it came from poor economic conditions in Venezuela, it makes sense to assume that few migrants have enough physical and material resources to justify the high risk of moving out if they have an informal job in the border cities. Therefore, only the young migrants willing to accept the risk of searching for better income opportunities might be interested in moving out, leaving behind their families until they arrive and obtain enough resources to bring their relatives in a second phase. In this regard, the success of the young migrants in achieving a sustainable income source at the final destination is critical for explaining the migration second phase. Modeling such movement in the second phase requires instrumental variables to control for the endogeneity caused by the success likelihood of young migrants achieving employment.

While instrumental variables modeling resolves endogeneity problems, the most noticeable difficulty in measuring the impact is the absence of a counterfactual outcome, i.e., the trend in wages, informal workforce, and occupation in border territories without the effect of migratory flows is unobservable (Callaway et al. 2018). Therefore, it is necessary to define a criterion to establish to what extent the observed change in the variables of interest results from the migratory flow or, on the contrary, is a consequence of natural growth in the informal workforce and the variables that influence them in a particular

context. In the reviewed literature, the method for constructing the control group used to identify the counterfactual of the treatment group considers control groups assembled from combinations of observations with similar characteristics to the observations of the treatment group, comparison with a previous and subsequent scenario in the panel data models with fixed effects for the same units of observation, and the definition of the control group using observations of people who behave similarly but have not faced the migratory flow (Callaway et al. 2018; Fan and Yu 2012; Roodman et al. 2019). Once a control group is available, comparing the trends of the treatment group and the control group in the variables of interest allows the assembly of the counterfactual of the treatment group, from which it is possible to determine the impact of the migratory flow. In the specialized literature, the parallel trends assumption, which identifies the counterfactual of the treatment group estimated from the control group, is a critical aspect of the econometric identification strategy.

Unlike the reviewed literature, this study estimates the impact of the first phase of the migratory flow on different definitions of informality without including the more endogenous second phase. In order to identify the impact on informality, the study constructed a control group with departments that met three conditions: (1) the department was not a border region, (2) the department had a similar development environment to the border departments in the control group, and (3) the department had a lower incidence of migratory flow from Venezuela. Thus, the control group comprised the Caldas, Choco, Huila, and Tolima departments. On the other hand, the border departments with Venezuela in the treatment group were La Guajira, Cesar, Norte de Santander, Arauca, Vichada, and Guainía. Boyacá, Cesar, Norte de Santander, and La Guajira departments have continuous and systematized labor market information because they are part of the Integrated Household Survey (IHS) sample implemented each quarter by the National Administrative Department of Statistics (NADS). On the other hand, the departments of Arauca, Vichada, and Guainía did not have continuous information during the research's interest period; therefore, the treatment group corresponding exclusively to departments bordering Venezuela was composed of Boyacá, Cesar, Norte de Santander, and La Guajira. Appendix A describes the criteria for selecting the control group and presents the indices of the departmental economic environment of the treatment group and the control group.

3. Migration Demographics

The section describes the characteristics of the migratory flow originating from Venezuela from 2013 to 2017. Table 1 presents the migratory flow by gender, age, and education. The participation of women in the migratory flow has progressively increased, reaching parity with men. In 2013, for every 100 people migrating from Venezuela to Colombia, there were 48 women and 52 men; by 2017, the participation was equal, with 50 men and 50 women per every 100 migrants. The turning point in female participation was in 2016, when male and female participation was equal. This trend is evident in their participation of the total migrants and their contribution to the variation, which progressively increased until reaching parity in 2017.

The age group with the lowest participation in the total migratory flow and the total annual variation was 65 years and older. Conversely, the age groups with the highest participation in the migratory flow were 0 to 18, 19 to 29, and 29 to 40. These groups also represent the highest contribution to the annual variation in migratory flow from 2014 to 2017; this implies a greater involvement of the young population in the migratory flow and a higher increase in the migrant population from 2013 to 2017. It is important to note that the 0 to 18 years group, considered a minor age group, represents a significant segment of the migratory flow and contributed the most to the variation in the total migratory flow.

Regarding the educational level, the groups with the highest annual participation were those without primary and intermediate education. These groups also provided the highest contribution to the annual variation; this implies a migratory flow characterized by unskilled labor, which has been increasing faster within the total migrant population since 2013. Those with the levels of higher education and postgraduate education maintained an average annual participation of 5% from 2013 to 2017, but their contribution to the variation is decreasing. On the other hand, the average years of education has remained below eight years, experiencing slight growth in 2017. In 2013, the average was 6.99 years; in 2017, the average schooling years reached 7.41 years. Therefore, the migratory flow from Venezuela was characterized by individuals aged 19 to 64. Additionally, migrant laborers, on average, had fewer than eight years of education, with a higher participation of groups with no education, primary education, and intermediate education.

Table 1. Migration demographics.

Variable	Demographics Group	2013	2014 Part.	2014 Cont. to % Var.	2015 Part.	2015 Cont. to % Var.	2016 Part.	2016 Cont. to % Var.	2017 Part.	2017 Cont. to % Var.
Gender	Male	52%	54%	61%	52%	49%	50%	48%	50%	50%
	Female	48%	46%	39%	48%	51%	50%	52%	50%	50%
Age	0 to 18 years	27%	34%	58%	37%	42%	18%	3%	38%	51%
	19 to 29 years	24%	25%	28%	23%	20%	24%	38%	26%	22%
	29 to 40 years	24%	22%	15%	22%	22%	22%	34%	22%	17%
	41 to 64 years	21%	17%	3%	17%	17%	15%	20%	13%	8%
	More than 65 years	5%	3%	−4%	1%	−2%	2%	4%	2%	2%
Education	Without education	30%	32%	41%	32%	40%	31%	32%	29%	29%
	5 years of education	31%	27%	13%	22%	18%	26%	31%	22%	19%
	9 years of education	3%	5%	13%	4%	3%	4%	4%	3%	2%
	11 years of education	18%	22%	28%	22%	27%	23%	25%	26%	33%
	13 years of education	3%	4%	8%	4%	5%	4%	4%	6%	10%
	16 years of education	5%	2%	−9%	3%	5%	4%	5%	5%	7%
	18 years of education	2%	1%	−3%	1%	1%	0%	1%	0%	0%
	Average years of education	6.99	6.63		6.82		6.96		7.41	
Total		34,680	44,407		75,634		166,473		291,441	

Source: own elaboration based on World Bank's (2018).

4. Methodology

4.1. Econometric Model

Measuring the impact of migration assumes the decision to migrate as exogenous and random in the first phase of the migration flow, affecting the bordering departments with Venezuela. Following the literature (Card 1990; Friedberg 2001; Tumen 2016) on migration published in the last ten years, the decision to migrate can be voluntary and endogenous when influenced by information about the labor market conditions in the host country. On the contrary, migration can be involuntary and exogenous when there is no correlation between the decision to migrate and the labor market conditions of the host country. In this specific case, massive and rapid migrations resulting from adverse economic conditions in Venezuela have been involuntary and exogenous, creating the required conditions of randomness in estimating the impact of migration on the host country's labor market. Therefore, the econometric model's identification strategy investigates the effect of migration on the labor market in its first phase using the difference-in-differences (DID) model, employing the panel data method with fixed effects to control for characteristics that influence the migration effect but remaining constant over time, such as a developmental variable associated with previous economic relationships between cities and departments

with Venezuela. Equation (1) describes the econometric model used to identify the fixed effects estimator of the impact of the migration flow on informality.

$$Inf_{i,t} = \beta_0 + \beta_1 TG_{i,t} + \beta_2 2016 - 2018_{i,t} + \beta_3 TG_{i,t} \times 2016 - 2018_{i,t}$$
$$+\beta_4 EDE_{i,t} + \beta_5 WE_{i,t} + C_i + D_t + \varepsilon_{i,t} \tag{1}$$

where $Inf_{i,t}$ is the number of workers in informal conditions for sample unit i, which could be departments, capital cities, or individuals, in the period t. $TG_{i,t}$ is a dichotomous variable that classifies sample units as belonging to the treatment group when equal to 1. According to the descriptive analysis presented in Section 3, 2016 was a breaking point in migration growth, with an increase of 137.57% compared with 2015. Thus, 2016–2018 is a dichotomous variable that takes the value of 1 to capture the behavior of the informal workforce after 2015, when the influx of migrant workers began to be massive, rapid, and involuntary.

$TG_{i,t} \times 2016 - 2018_{i,t}$ is an interaction variable between the treatment group and the observation window. The control variables in measuring the impact were the education of employed individuals ($EDE_{i,t}$) and the work experience of employed individuals ($WE_{i,t}$). C_i corresponds to the fixed effects of each sample unit i, and D_t indicates the time constant of each sample unit i. $\varepsilon_{i,t}$ represents the estimation errors for observation unit i in each period t. To estimate the probability of change in informality outcomes in the treatment group, we used a probabilistic model with a normal distribution known in the literature as the Probit model. Rosales et al. (2013) and Cameron and Trivedi (2009) provided the theoretical and applied discussion of the Probit model.

Education is considered a fundamental investment in human capital. When individuals invest in their education by acquiring skills, knowledge, and competencies, they increase their human capital, which translates into higher workplace productivity and income over time. Thus, the level of education is related to the number of years of formal education an individual has acquired. The higher the level of education, the more skills and knowledge individuals have, which in turn influences their ability to access higher-paying jobs and positions of greater responsibility. Work experience is also considered an investment in human capital. As individuals gain experience in the workforce, they develop specific skills and knowledge related to their fields. Furthermore, the time a person has worked in a particular industry or position is critical in providing individuals with specific knowledge and skills to perform better at work. Accumulated work experience over the years is often associated with improved skills and efficiency in the workplace.

Specific dynamics affecting individuals can influence how variables related to education and work experience impact the lives and job opportunities of migrants and native workers. From the perspective of migrant workers, such influence includes a geographical transition and often a change of the work environment. In this regard, migrant workers' education and work experience may or may not be directly transferable to the new destination. Moreover, they may have to pursue additional educational requirements to get the recognition of general and specific human capital accumulation, qualifications, and work permits for better integration into the formal job market segment. Therefore, migrants face employment and educational barriers that may block them from participating in formal labor markets, ending up in informal segments with lower educational and work experience requirements.

From the perspective of native workers, such specific dynamics relate to how the arrival of migrant workers affects their employment opportunities. In the context of an unregulated migration with employment barriers to the most qualified migrant workers, native workers with a lower level of human capital may experience more difficulties accessing jobs in the informal segment of the labor market due to an increase in the quality of labor supply available for open vacancies at a given level of informal production. Moreover, when a migrant worker does have the same level of human capital accumulation but faces barriers to obtaining a work permit, the wage differentials will displace native workers with a lower level of human capital accumulation. Therefore, human capital

accumulation offers native workers better possibilities to overcome an increase in labor supply when there is a lack of migration policies that provide instruments for promoting complementary technologies between the human capital accumulation of both migrant and native workers. Thus, education level and work experience control the impact of the migration flow on the native workers. Thus, the years of education and work experience of native workers influence the impact of the migration flows on native workers' informality levels. In addition, if the incoming migrant workers expand the informal production, there will be an effect experienced through increases in labor productivity, as better conditions are provided to increase native job opportunities in the informal labor market segment. Thus, wages remain without significant changes because an increase in labor demand offsets the labor supply increments due to the migration flow.

A 2013 to 2018 departmental database with a pre-impact period between 2013 and 2015 and a post-impact period from 2016 to 2018 established the time framework during which the migration impacted the natives' informal workforce outcomes. Furthermore, the pre-impact period and the post-impact period in the city and individual databases were from the fourth quarter of 2015 to the fourth quarter of 2018 and from July 2015 to December 2018, respectively.

The econometric model aims to identify the effect of forced, massive, and rapid migration from Venezuela on the native informal workforce from 2016 to 2018, with 2013 to 2015 as an ex ante period and from 2016 to 2018 as an ex post period. The research hypothesis is defined as follows:

H$_1$: *Forced migration from Venezuela increases the native informal workforce in Colombia.*

Considering the research hypothesis and the identification of the econometric model, the coefficient of interest in Equation (1) is β_3, which reflects the impact on the informal workforce in the departments bordering Venezuela based on constructing a counterfactual represented by the departments selected in the control group. Appendix A presents the selection criteria for the control group. The expected sign of the coefficients in the three subsamples is positive.

Figure 1 describes the estimation strategy to identify the impact of the migration flow on informality. The fixed effects estimator of the effect measures the difference in informality averages between the control group and the treatment group over time, controlling for education, work experience, and fixed effects of sample units over time. For example, for two time periods, T = 1 and T = 2, the informality results in the treatment group $(Y_1|D = 1)$ are isolated by constructing an informality trend without migration flow using the informality results in the control group $(Y_1|D = 0)$. A difference model captures the difference in informality results between the treatment group and the control group (β_D). In contrast, a difference-in-differences model captures the effect of the migration flow by comparing the current informality trend in the treatment group and the counterfactual trend of what would have happened without the migration flow, using the observed trend in the control group $(\beta_{DID} = \beta_3)$. The estimation requires the assumption of parallel trends in informality behavior for the control and treatment groups and that there is no correlation between the migration decision and the labor market characteristics of the host departments of the migration flow.
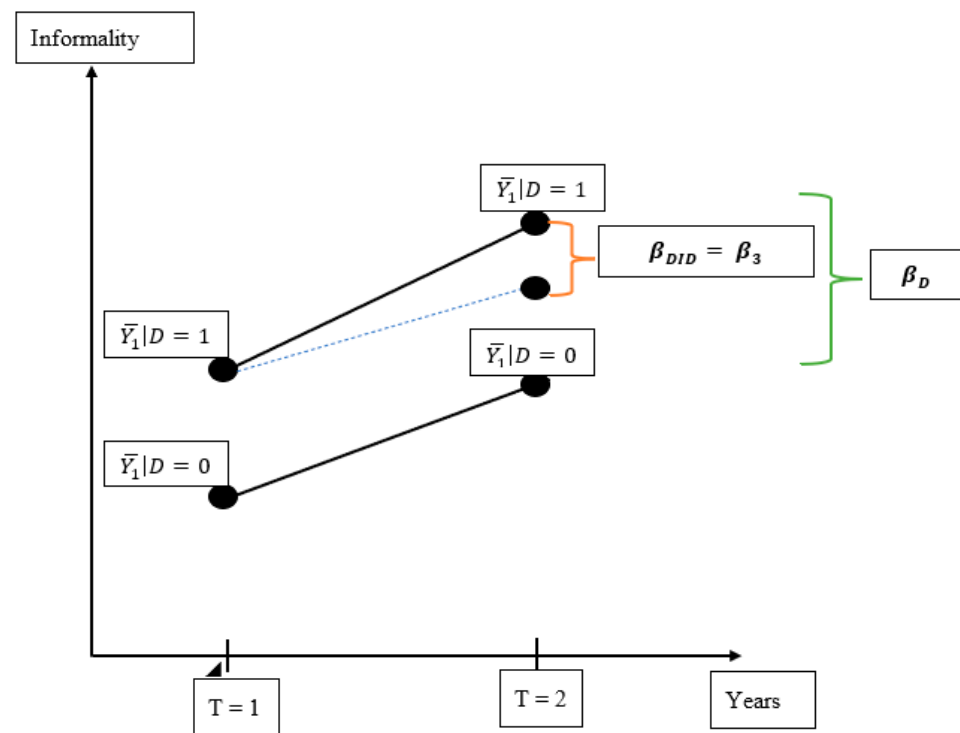
Figure 1. Strategy for estimating the effect of migration on informality. Own elaboration based on Bernal and Peña (2011).

4.2. Data

The source of information for constructing informality-related variables was the Integrated Household Survey (IHS), which includes a quarterly representative sample for individuals and capital cities and an annual representative sample for departments assessing labor market performance. The IHS is a survey focused on individuals. Thus, the information obtained reflects the perspective of native workers regarding their employment status, labor income, informality conditions, occupations, and occupation sectors, among other labor aspects. The HIS's microdata were publicly accessible on the National Administrative Department of Statistics (NADS)'s website in Colombia. Subsequently, a database was constructed, including variables related to the informal workforce indicators, education, and work experience for the control and treatment groups.

4.3. Definition and Interpretation of Variables

Table 2 describes the variables constructed to measure the impact of rapid and massive migration in the departments bordering Venezuela from 2013 to 2018 on three different informal workforce variables. The first variable relates to the informal workforce defined according to NADS criteria, primarily related to company size, inclusion of unpaid workers, and self-employment of workers. The second variable pertains to employed individuals who report not making social security payments associated with health and pensions in the household survey. The third variable represents the percentage participation of the informal workforce in the total employed population. This variable measures the logarithmic gap between the informal workforce and total employment. A higher percentage of participation in informality indicates a more significant gap between the informal workforce and employment, resulting in reduced social protection for workers.

Table 2. Informality definitions.

Informality	Description	Definition
Informality based on the firm's size	Number of workers in informal conditions.	The NADS's criteria for determining the informal workforce size correspond to (1) private employees and workers who work in establishments, businesses, or companies that employ up to five people in all agencies and branches, including the employer and/or partner; (2) unpaid family workers; (3) unpaid workers in enterprises or businesses of other households; (4) domestic employees; (5) day laborers or laborers; (6) self-employed persons working in establishments of up to five persons, except self-employed professionals; (7) employers in enterprises with five or fewer workers; (8) government workers or employees are excluded.
Informality based on social security payments	Number of workers in informal conditions.	Number of employed persons declaring in the household survey that they do not pay into social security. Social security payments include health and pension payments.
Informality as a share of employment	Number of workers in informal conditions/Number of employed persons.	Percentage share of informality in total employment. It measures the gap between employed and informal employees by criteria of the NADS and by criterion of payment of social security contributions. In this sense, two gap variables are constructed to fit the econometric model.
Education (X_1)	Average years of schooling of employed people.	Average characteristics of departments and cities and individual characteristics of employed persons as control variables that mediate the impact of migratory flow on informality and labor income. Greater education and work experience of the employed imply greater barriers to access to the labor market by migrants.
Work experience (X_2)	Proxy of the average work experience for years employed (Age—15 years—years in education).	

Source: own elaboration based on the IHS.

The HIS's data are the source for the control and the treatment group variables. The education averages for departments and cities aggregate individual data registered at each geographical sample. Moreover, in the case of the individual subsample, education was measured based on the number of years of education reported in individual IHS responses.

The work experience variable is synthetic and constructed from three IHS variables: (1) age, (2) age to participate in the labor market, and (3) years of education. The estimation of this variable partially captures the potential exposure of different individuals in the subsamples to the labor market. Both control variables for the quantity of education and nominal work experience are measured in years, without reference to education quality or specific experience for each individual, as the household survey does not collect specific information on education quality, and the individual sample changes from quarter to quarter. The pseudo-panel characteristics of the household survey pose challenges in longitudinally tracking sample units for departments, capital cities, and individuals because it interviews a different sample over time, but it changes the sample each quarter. Moreover, the selection of control variables aimed for parsimony, guided by a literature review in the identification of the econometric model, to determine the extent to which the estimation of the impact of migration was affected by the selected control variables.

The sample of departments is representative of the annual frequency. Hence, the control group has 24 observations (6 years by four departments), and the treatment group has 24 observations, for a maximum total of 48 observations. The observation window starts from 2013 to 2015, the period before the informal workforce trend change, and ends with 2016 to 2018, considered the period after the informality trend change. Regarding capital cities data, which are representative at the quarterly level, the control group has 96 observations (6 years by four cities by four quarters), and the treatment group has 96 observations, for a maximum total of 192 observations. The observation window starts in the first quarter of 2013 and ends in the last quarter of 2018.

Finally, the sample of individuals has a maximum of 2,645,065 people distributed into 1,291,586 people in the treatment group (48%) and 1,353,479 people in the control group (52%). The number of observations may vary depending on the availability of information to estimate control variables and informality variables for the sample unit. Unlike the related literature, where changing the treatment group allows sensitivity analysis of the migration effect on the informal workforce, this research incorporates different subsamples at the departmental, metropolitan, and individual levels to investigate significant changes in impact assessment and its corresponding statistical significance.

5. Descriptive Analysis of Informality Trends

From 2013 to 2015, the behavior of informality, on average, for each group showed a similar trend for the treatment group and the control group in departments, capital cities, and the sample of individuals. However, the number of native workers in informal conditions increased from 2016 in all subsamples, with a faster pace in the treatment group's average compared to the control group. The trend's break coincides with the highest percentage increase in migrants born in Venezuela, observed between 2015 and 2014.

Table 3 describes the trends in size and characteristics of the informal workforce in the control group (CG) and the treatment group (TG), along with variables related to the employed's education and years of work experience in the sample of departments. Informality followed a similar trend from 2013 to 2015 in the sample of departments. However, the level changed in 2016, with an increase in informality among native workers of 545,050 people in the control group and an increase of 877,861 people in the treatment group. The variables measuring educational background and work experience follow the same trend, with a level change in years of education starting in 2016. Work experience shows slight changes since 2016, with a reduction, particularly for the treatment group.

Table 3. Departments: characteristics of the native workforce.

Variable	2013	2014	2015	2016	2017	2018
Informal workers_CG	234,327	371,225	351,123	896,173	898,564	897,771
Informal workers_TG	261,643	366,688	394,970	1,272,831	1,277,704	1,255,522
Employed's years of schooling_GC	7.24	6.94	7.14	6.51	6.57	6.71
Employed's years of schooling_TG	7.3	7.26	7.26	6.91	6.97	7.05
Work experience_CG	21.32	22.35	21.7	22.08	22.1	22.05
Work experience_TG	22.92	22.92	23.16	21.6	21.7	21.9

Source: own estimate based on IHS data. CG is control group and TG is treatment group.

Figure 2 depicts the behavior of informality trends at the departmental level. As observed, the departmental informal workforce trends for both the control and treatment groups follow a parallel trajectory. However, starting in 2015, the informality growth rate in the treatment group accelerated faster than in the control group. Between 2015 and 2016, informality in the control group increased by 155%, whereas during the same period, informality in the treatment group exhibited a growth rate of 222%. Post-2016, informality stabilized with positive but higher growth rates in the treatment group. Between 2017 and 2016, informality increased in the control group at a rate of 0.27%, while in the treatment group, it grew by 0.38%. Towards the end of 2018, the growth rate decelerated, with a decrease of 0.09% in the control group and 1.74% in the treatment group. When examining the absolute values in Table 2 in conjunction with Figure 2, it is evident that the change observed in 2015 modified the average informality rate, placing informality higher than that observed in 2014, mainly for border departments.
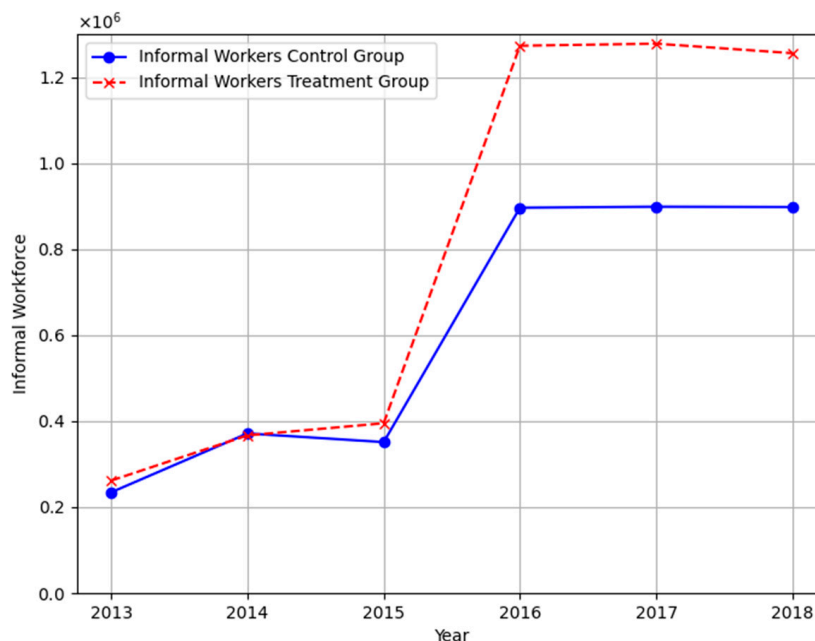
Figure 2. Parallel trends in workers' informality in the sample of departments.

Table 4 presents the behavior of informality in the capital cities of the control and treatment groups. Similarly to the departmental sample case, the informality level in the capital cities changed starting in 2016. In the control group, informality in 2016 increased to 271,285 individuals, which corresponds to an increase of 161,987 individuals compared to 2015 and represents the highest informality growth from 2013 to 2018. On the other hand, in the treatment group, there was an increase of 325,603 individuals from 2015 to 2016. Similarly, this increase in informality levels is the highest in observed growth from 2013 to 2018. The observed growth influenced the long-term informality trend, as the levels observed in 2017 and 2018 are above those observed in 2014. Furthermore, the variables measuring education and work experience characteristics remain very similar throughout the time sample, with a slight decrease in work experience starting in 2016, both in the control and treatment groups.

Table 4. Capital cities: characteristics of native labor.

Variable	2013	2014	2015	2016	2017	2018
Informal workers_CG	111,419	112,166	110,398	272,385	271,153	272,163
Informal workers_TG	157,487	165,808	164,850	490,453	491,761	484,796
Employed's years of schooling_GC	8.5	8.8	9.01	8.37	8.4	8.5
Employed's years of schooling_TG	8.8	8.8	8.97	8.57	8.64	8.74
Work experience_CG	20.3	20.5	20.2	19.2	19.1	19.1
Work experience_TG	20.8	21.12	21.5	19.1	19.2	19.1

Source: own elaboration based on the IHS data. CG is control group and TG is treatment group.

Figure 3 illustrates the growth rate of informality in capital cities for both the control and treatment groups. The behavior of informality in both groups underwent a significant change starting in 2015. From 2015 to 2016, the informality growth rate was 147% in the control group and 198% in the treatment group. Subsequently, the informality growth rate stabilized in both groups, with a more pronounced change in the absolute values determining the level of informality in the treatment group compared to the control group.
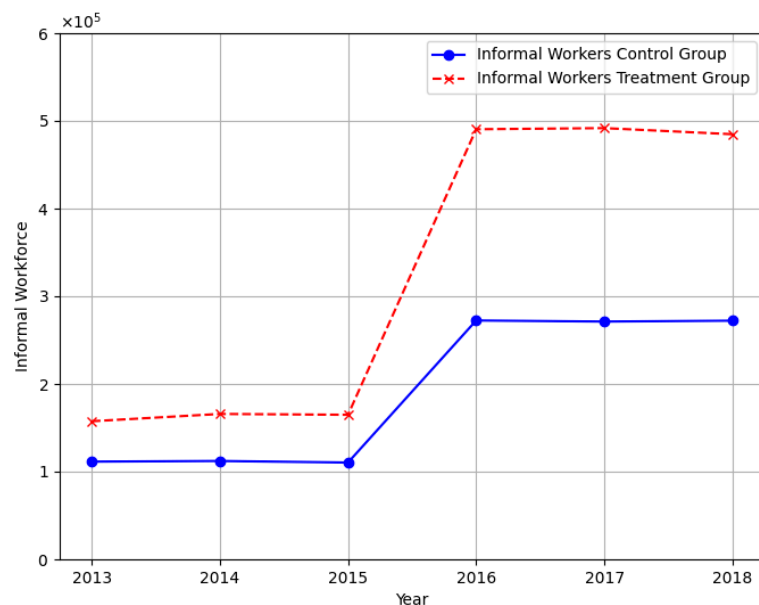
Figure 3. Parallel trends in workers' informality in the sample of capital cities.

Table 5 describes the characteristics of the sample of individuals in the control and treatment groups. In both groups, the number of people in informal conditions notably increased from 2015 to 2016. In the control group, informality in 2016 reached 100,700 people, implying an increase of 61,292 people compared to 2015. For the treatment group, the behavior was similar, with an informality level of 109,339 in 2016, corresponding to a growth of 71,546 people compared to 2015.

Table 5. Individuals: characteristics of the native workforce.

Variable	2013	2014	2015	2016	2017	2018
Informal workers_CG	28,869	40,368	39,408	100,700	98,954	94,996
Informal workers_TG	27,793	37,617	37,793	109,339	107,515	102,509
Employed's years of schooling_GC	8.5	8.42	8.55	8.01	8.06	8.23
Employed's years of schooling_TG	8.3	8.32	8.41	8.04	8.11	8.21
Work experience_CG	20.3	20.5	20.2	19.2	19.1	19.1
Work experience_TG	20.8	21.12	21.5	19.1	19.2	19.1

Source: own elaboration based on IHS data.

Before 2015, the behavior of informality for both groups was similar, and the level remained close to 100 people after 2015. The informality levels in 2017 and 2018 were higher for the treatment group than those observed in the control group. The control variables measuring educational attainment and work experience remained unchanged from 2013 to 2018 in both the control and treatment groups. In general, the average years of schooling remained at eight years in both groups, and work experience experienced a slight decrease from a level of 20 years starting in 2015, averaging 19 years.

Figure 4 depicts the growth rate of informality for individuals in the sample of the control group and the treatment group. Although the trends are parallel, there is a higher increase in the treatment group's informal workforce. In particular, the growth rate shifted from 2016 onwards, with a 156% increase in informality among individuals in the control group and a 189% increase among individuals in the treatment group.
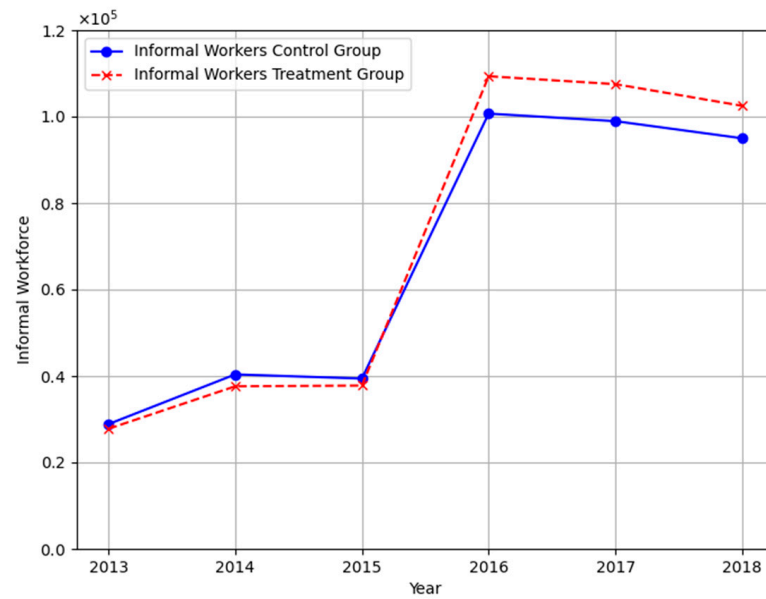
Figure 4. Parallel trends in workers informality in the sample of individuals.

Departments, capital cities, and individuals aligned with the change in the migration flow trend. Figure 5 depicts the behavior of informality among individuals who were part of the treatment group and the total migration flow, the migration flow of individuals born in Venezuela, and the migration flow of individuals born in Colombia returning to the country from 2013 to 2017. The graph shows the informal workforce trend within the treatment group and migration change trends in 2015. While in 2016, the informal workforce within the treatment group increased by 189% compared to 2015, the total migration grew by 120% in the same year, with a more significant increase in migration of individuals born in Venezuela (207%) compared to the migration of returned individuals born in Colombia (71%).
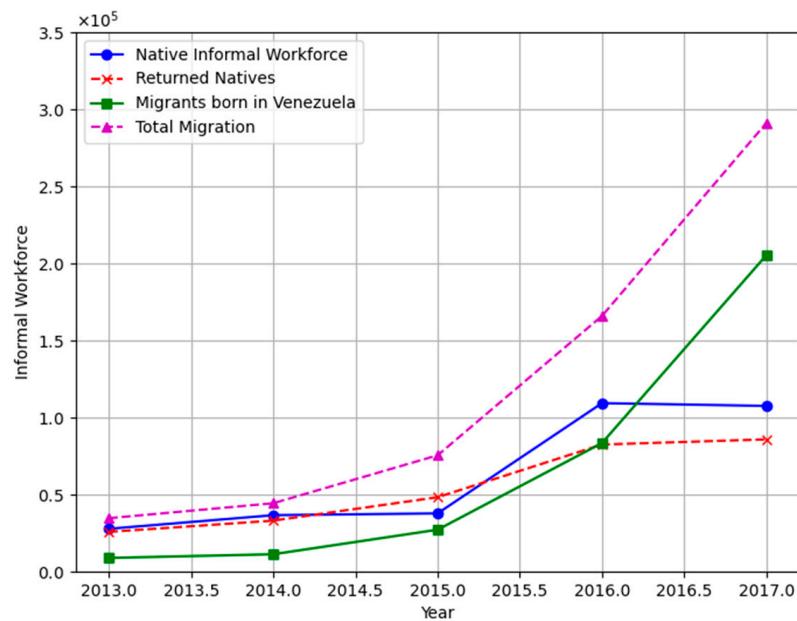

Figure 5. Informality trend for 2013–2018 in the sample of individuals.

6. Results

6.1. Effects on the Informal Workforce Based on the Firm's Size

The first and second columns of Table 6 contain the estimation results for the departmental subsample without control variables (column 1) and with control variables (column 2). Comparing columns 1 and 2 allows us to determine the extent to which the effect of migration depends on the control variables in terms of effect size and statistical significance. Similarly, columns 3 and 4 present the estimates made for the capital city subsample. Column 3 describes the adjustment without control variables, while column 4 presents the adjustment with control variables. Finally, columns 5 and 6 make the same comparison for the individual-level adjustment.

Table 6. Effects on informal workforce based on the firm's size.

Variable	Departmental Model	Departmental Model with Control Variables	Municipal Model	Municipal Model with Control Variables	Individual Model	Individual Model with Control Variables
Effect of migration	108,545 **	107,933 *	43,922 ***	35,139 ***	−0.027 ***	0.025 ***
Time	12,357	12985	1035 *	1420 **		
Education		−3417		−18,857		−0.044 ***
Work experience		−10,177		−17,796 ***		0.0036 ***
Constant	319,000 ***	258,051	42,841 ***	607,843 ***	0.5723 ***	0.89 ***
N	48	48	184	184	3,157,762	1,344,352
R^2 and pseudo-R^2	0.61	0.60	0.64	0.68	0.00	0.1607
F and Chi-Square	72.78	41.34	102.95	89.28	149.01	290,832

Source: own elaboration based on data from the IHS. The sample of departments in the treatment group is restricted to departments of Boyacá, Cesar, Norte de Santander, and La Guajira, bordering Venezuela, excluding Arauca, Vichada, and Guainía. The sample of capital cities in the treatment group corresponds to Tunja, Valledupar, Cúcuta, and Riohacha. The control variables used were average years of schooling and a proxy variable of work experience for those employed. *** Significance Level of 1%. ** Significance Level of 5%. * Significance Level of 10%. The individual adjustment was performed with a Probit model to estimate the marginal effect of a change from the dichotomous variable independent of the state of non-formality to the state of formality. The definition of informality used corresponds to the condition of those employed in companies of up to five people.

The rows of Table 6 present the estimation of the migration's effect on informality in terms of the number of people, the natural growth inertia of informality (time), control variables related to education and work experience, and the estimation constant in a model with fixed effects. Subsequently, the table reports the observations for each adjustment in the rows, two goodness-of-fit measures (R^2 and pseudo-R^2), and two tests of the joint hypothesis of all coefficients being equal to zero (F-test and Chi-Square test).

In columns 1 and 2, the migration effect positively affected the departmental adjustment of the informal workforce. Additionally, the coefficient size remains relatively unchanged with the inclusion of control variables and is statistically significant at 5% without control variables and at 10% with control variables. Considering the result from column 2, migration's effect on border departments increased informality by 107,933 individuals. The natural or inertial growth in informality corresponds to 12,985 individuals per year. Similarly, one-year increases in education reduced informality by 3417 individuals, while one-year increases in work experience reduced informality by 10,177. Although the expected signs of the control variables correspond to the anticipated relationship, the coefficients are not statistically significant, and their inclusion is solely justified to assess changes and stability in the migration effect. The model with control variables explains 60% of the variance in informality (R^2), and the null hypothesis that all model coefficients are simultaneously equal to zero is rejected based on the results of the F-test.

Likewise, the migration effect on capital cities' adjustment is positive without control variables (column 3) and with control variables (column 4). The estimated effect of migration is significant at 1% in the model without control variables and in the model with control variables. The effect size is smaller in the model with control variables, resulting in an increase of 35,139 individuals in informal conditions due to migratory flow, compared to an increase of 43,922 individuals in informal conditions in the model without

control variables. The natural or inertial trend of informality corresponds to an increase of 1035 individuals per quarter. The control variables of education and work experience have the expected negative sign, with reductions in the informal workforce of 18,857 and 17,706 individuals, respectively, for each additional year of education and work experience. Unlike the departmental model, the estimated coefficients of the control variables are significant for the inertial growth in informality and work experience. The improvement in statistical significance is associated with increased observations, which reduces the standard deviation compared to the departmental adjustment. It also enhances the explanatory capacity of informality variance, increasing from 60% in the departmental adjustment to 68% in the capital cities' adjustment, as measured by R^2. The joint test of the model variables' relevance to explain informality (F-test) rejects the null hypothesis that all model coefficients are statistically equal to zero.

In the individual adjustment presented in columns 4 and 5, the control variables change the direction of the migration effect on informality. Without controlling for education and work experience, the model estimates a 2.7-percentage-point decrease in the probability of informality among native workers, significant at 1%; this would result in the substitution of native informal workers by migrant informal workers while keeping the informal labor market segment constant. However, when controlling for education and work experience, the effect of migration on the informality of native workers is positive, with a 2.5-percentage-point increase in the probability of informality, significant at 1%. Thus, human resources are incorporated into the informal segment, increasing informality frequency.

Furthermore, an additional year of education reduces the probability of informality by 4.4 percentage points, significant at 1%. At the same time, work experience has a null effect on the probability of informality, though it is statistically significant at 1%. The probabilistic model requires a transformation with marginal effects to determine changes in percentage points in the probability of being in the informal workforce. Table 6 presents the estimation results. Pseudo-R^2 estimates allow for comparing the model without control variables and the model with control variables to determine which model more accurately predicts informality occurrence, with a binary variable in the individual model equal to 1 when the individual is in informality. The model with control variables has a better predictability of informality, as indicated by the pseudo-R^2 value (0.16). The hypothesis test conducted using the Chi-Square distribution suggests that the included variables in the model are relevant for explaining informality.

6.2. Effects on the Informal Workforce Based on Social Security Payments

Table 7 presents the results of the estimations for the informal workforce measured as contributions to the social security system. The effect of migratory flow on informality is positive for all subsamples without control variables and is statistically significant at 1%. When including control variables, the positive direction of the effect remains, as well as statistical significance, only in the departmental and capital cities subsample. In the individual subsample, the effect of migration changes signs and is not statistically significant. However, education and work experience increases are statistically significant, with adverse and null effects on the probability of being in the informal workforce, respectively. Such results may be associated with heterogeneous effects on the informality distribution when controlling for education and work experience. Castillo et al. (2022) describe the impact on salary distribution, pointing to heterogeneous effects in salary quartiles.

From the previous finding, the average effect might not represent migration because substitution could occur for some informal workforce groups with different education and work experience characteristics. Migratory flow may be affecting different informal groups in distinct ways. Considering that the years of education for members of the migratory flow are similar to those of the native population (8 years) in some informal labor market segments, the expulsion of native informal workers may occur in certain occupations. However, as indicated by the results, estimates depend on the sample, so there is no statistical significance to infer that, on average, there is a substitution of native

by migrant workers in the informal segment of the labor market. Thus, this implies a benefit to modeling the effect on the distribution of informal occupations to determine if the migration effect is more significant for certain native informal workers in terms of the expulsion of native informal employees or expansion of the informal segment, controlling for education and years of work experience.

Table 7. Effects on informal workforce as contributions to the social security system.

Variable	Departmental Model	Departmental Model with Control Variables	Municipal Model	Municipal Model with Control Variables	Individual Model	Individual Model with Control Variables
Effect of migration	348,975 ***	361,966 ***	153,327 ***	118,696 ***	0.284 ***	−0.0035
Time	29,301	28,801	4849 **	6227 ***		
Education		−8886		−68,533 *		−0.0483 ***
Work experience		5867		−68,488 ***		0.003 ***
Constant	271,392 ***	213,766	107,737 ***	224,174 ***	0.085 ***	1.497 ***
N	48	48	184	184	3,157,762	1,344,352
R^2 and pseudo-R^2	0.72	0.72	0.65	0.74	0.0014	0.1823
F and Chi-Square	105.67	60.17	108.76	98.50	4061	322,420

Source: own elaboration based on data from the IHS. The sample of departments in the treatment group is restricted to departments of Boyacá, Cesar, Norte de Santander, and La Guajira, bordering Venezuela, excluding Arauca, Vichada, and Guainía. The sample of capital cities in the treatment group corresponds to Tunja, Valledupar, Cúcuta, and Riohacha. The control variables used were average years of schooling and a proxy variable of work experience for those employed. *** Significance Level of 1%. ** Significance Level of 5%. * Significance Level of 10%. The individual adjustment was performed with a Probit model to estimate the marginal effect of a change from the dichotomous variable independent of the state of non-formality to the state of formality. The definition of informality used corresponds to social security payments made by the employed.

Regarding the magnitude of the migration effect, the estimation results suggest an increase of 361,966 individuals in the informal workforce at the departmental level and 118,696 individuals at the metropolitan level in the adjustments with control variables. In the individual adjustment without control variables, the probability of informality increases by 2.8 percentage points. In contrast, adjusting with control variables offsets the effect of being in the informal workforce. Thus, there is a potential for heterogeneous effects on the informality distribution. Linear adjustments with control variables for departments and capital cities explain 72% and 74% of the informality variance, while the non-linear model with control variables better predicts the probability of informality with a pseudo-R^2 of 0.18. F-tests for the linear models in the department and capital city subsamples and the Chi-Square test for the individual model indicate the relevance of the included variables in their respective adjustments for explaining informality.

Research findings suggest a positive response in informal jobs for native workers without a significant change in the number of native informal workers related to total employment. Following the economic theory, the research findings seem to follow a transmission mechanism in which migrant workers reduce labor costs and increase production in informal markets, providing better conditions to increase informal jobs for native workers. Hiring more qualified or equally qualified informal migrant workers at a lower wage increases the opportunities to hire native workers because they increase labor productivity over labor costs at a given level of demand for informal production. Such an improvement in informal production technology causes an expansion in the informal segment that provides a case for showing the effects on labor productivity as a primary driver of an increase in informal job opportunities for native workers. There is, however, a role for education and work experience in the individual model that might indicate heterogeneous effects on the informal workforce distribution along the years of education and work experience.

6.3. Effects on the Informal Workforce as a Share of Total Employment

Table 8 describes the effect of migration on the relative size of informality according to criteria from NADS concerning employment in firms with at most five employees. In the departmental adjustment, the estimates made regarding informality participation

reflect a decrease in the adjustment without control variables and the adjustment with control variables. The informal workforce participation in the adjustment with control variables decreases by two percentage points. However, the effect is insignificant due to the limited number of observations, which increases the standard deviation and decreases the probability of rejecting the null hypothesis that the coefficient equals zero. In contrast, in the municipal adjustment with more observations, the relative size of informality increases by one percentage point in both the model without control variables and the model with control variables. Education and work experience reduce the relative size of informality but are not statistically significant. Therefore, it is not possible to infer that in any sample of the control group and the treatment group, the effect of migration implies an expansion in the informal segment of total occupation.

Table 8. Effects on the gap between the informal workforce and the employed according to the definition of informality using employed persons by company size.

Variable	Departmental Model	Departmental Model with Control Variables	Municipal Model	Municipal Model with Control Variables
Effect of migration	−0.0039	−0.0207	0.0150	0.0194
Time	0.0018	0.0112	−0.0011	−0.0003
Education		−0.1427		−0.0322
Work experience		−0.0533		−0.0032
Constant	0.60 ***	3.005	0.5102 ***	0.8985
N	48	48	184	184
R^2	0.22	0.21	0.19	0.20
F	4.29	2.86	14.19	8.67
Degrees of freedom	37	35	173	171

Source: own elaboration based on data from the IHS. The sample of departments in the treatment group is restricted to departments of Boyacá, Cesar, Norte de Santander, and La Guajira, bordering Venezuela, excluding Arauca, Vichada, and Guainía. The sample of capital cities in the treatment group corresponds to Tunja, Valledupar, Cúcuta, and Riohacha. The control variables used were average years of schooling and a proxy variable of work experience for those employed. *** Significance Level of 1%. The definition of gap is the quotient between informal employment and total employment, where informal employment corresponds to NADS's informality criteria. The individual adjustment was not made because the independent variable counts the number of people in informality and the number of employed people possible only for the subsamples of departments and capital cities. The participation of informal workers relative to those employed reflects the relative size of informality by NDAS informality criteria.

A factor that could explain why the relative size of the informal workforce in employment does not change is the number of Colombian nationals that returned to Colombia after living in Venezuela, which increased employment figures. According to migration data, the number of Colombian natives' returning after living five years in Venezuela increased by 25% between 2014 and 2013, 35% between 2015 and 2014, 99% between 2016 and 2015, and 19% between 2017 and 2016. At the same time, the employment growth of native workers grew by 40% between 2014 and 2013, 1% between 2015 and 2014, and 95% between 2016 and 2015, and decreased by 2% between 2017 and 2016. Thus, the Colombian returned nationals influence employment because they do not face employment barriers related to documentation to be eligible for job openings in Colombia.

In comparison, as per the results reported in Table 9, the impact on the relative size of informality adjusted with the definition of informality based on social security contributions coincides with an increase in the participation of the informal segment in both the departmental adjustment and the capital city adjustment models, both with and without control variables. In the departmental adjustment, the effect of migration on informal participation is statistically significant at the 10% level. While the negative sign of education's impact on the size of informality is present in both the departmental and capital city models, work experience harms informal participation only in the capital city model.

The explanatory power of the variance in relative informality participation (R^2) is higher for adjustments based on social security contributions compared to models using informality criteria from NADS. Similarly, the probability of rejecting the null hypothesis of F-tests, which posits that the estimated coefficients included in the adjustment are different from zero, is higher in the adjustment using informality defined by social security contributions.

Table 9. Effects on the ratio of the informal workforce to employed according to the definition using contributions to the social security system.

Variable	Departmental Model	Departmental Model with Control Variables	Municipal Model	Municipal Model with Control Variables
Effect of migration	0.0932	0.1431 *	0.0685	0.0087
Time	0.0151	0.0150	0.0096 **	0.0063 *
Education		−0.0665		−0.1227 *
Work experience		0.0129		−0.0485 **
Constant	16855 ***	19544	14100 ***	12105
N	48	48	184	184
R^2	0.8350	0.8497	0.60906	0.7410
F	61.41	39.58	128.71	97.83
Degrees of freedom	37	35	173	171

Source: own elaboration based on data from the IHS. The sample of departments in the treatment group is restricted to departments of Boyacá, Cesar, Norte de Santander, and La Guajira, bordering Venezuela, excluding Arauca, Vichada, and Guainía. The sample of capital cities in the treatment group corresponds to Tunja, Valledupar, Cúcuta, and Riohacha. The control variables used were average years of schooling and a proxy variable of work experience for those employed. *** Significance Level of 1%. ** Significance Level of 5%. * Significance Level of 10%. The definition of gap is the quotient between informal employment and total occupation, where informal employment corresponds to NASD's informality criteria. The individual adjustment was not made because the independent variable counts the number of people in informality and the number of employed people possible only for the subsamples of departments and capital cities. The participation of informal workers relative to those employed reflects the relative size of informality by contributions to the social security system.

7. Discussion

Unlike the recent literature on migration's impacts in Latin America, which focuses on the second phase of migration to cities with more significant employment and income generation opportunities, this paper emphasizes the impact of migration on informality in labor markets in bordering departments, taking into account their economic development challenges. Therefore, this study's primary contribution lies in applying a methodology to measure the initial impact of migratory flow on informality in bordering departments, which initially face constraints in complementing the migrant flow.

The research adopts a quantitative approach, investigating correlations and causality between the growth in migrant numbers and the capacity to absorb labor in the informal segment within contexts characterized by resource access difficulties, institutional constraints, and the absence of labor migration policies to manage and structure the migration flow toward strategies that enhance sectorial productivity using microdata from the IHS (Integrated Household Survey) with variables operationalizing the theoretical concepts identified in the critical literature review. The critical literature review covered publications in indexed journals related to migration's impact over the last ten years.

The research design involved identifying a control group of departments not bordering Venezuela, with similar characteristics in terms of economic development context and informality, which did not experience the first migration phase due to their non-bordering status with Venezuela. The control group was used to construct the informality trend in bordering departments in the absence of migration, subsequently comparing this counterfactual with observed informality behavior and isolating the migration's effect on expanding the informal segment in bordering departments. This design, well-known in specialized litera-

ture as the difference-in-differences model, is frequently used in impact evaluations. The estimation included subsamples at the departmental, metropolitan, and individual levels to explore the sensitivity of the migration's impact parameter on informality estimated using different data samples. Additionally, the econometric model uses three distinct definitions of the informal workforce to investigate the parameter's sensitivity to the selected informality definition. The informal workforce definitions included occupation by company size and by social security contribution payments. Furthermore, a measure of the relative participation of informality, as the ratio between informal and total employment, was included to assess the extent to which migration has affected the relative size of informality.

The estimation results of the migration's effect on the informal labor market segment in bordering departments indicate an increase in the number of individuals in informal conditions by company size and social security contributions. However, the estimates were inconclusive regarding an increase in the relative participation of the informal workforce in total employment. Nonetheless, there is a positive effect of migration on the informal workforce size when using the informal workforce variable based on the social security definition.

The finding that migratory flow does not affect the informal-to-employed gap suggests redistributive effects in which an increase in employed individuals counteracts the informal workforce increment. In order to isolate these changes, it is necessary to consider the effect of migratory flow on the distribution of informal and employed individuals. In this sense, the absence of an effect on the percentage participation of the informal workforce in employment also suggests an effect on the distribution relevant to determining how much of the effect is permanent, transitory, or distributive. For example, some native workers may be taking jobs in the public sector where there are entry barriers for migrant workers, canceling out the effect on the informal-to-employed gap or the relative participation of informality. Moreover, a factor that could explain why the relative size of the informal workforce in employment does not change is the number of Colombian nationals that returned to Colombia after living in Venezuela, which increased employment figures.

Furthermore, estimations regarding the impact of migratory flow on the probability of informality indicated increases in the likelihood of informality occurrence in the definitions based on company size, with the magnitude and sign-changing with the inclusion of education and work experience. Including education and work experience indicates adverse changes in the likelihood of being in the informal workforce when using social security contribution payments to measure the informal workforce. Thus, the findings reflect heterogeneous effects on the distribution of informality at the individual level for each education and work experience group, making the average effect of the total impact statistically insignificant. Considering this, research on the impact of migration on the distribution of informality and the identification of effects on labor productivity is of interest to investigate the extent to which migration in bordering departments presents an opportunity for regional economic development.

8. Conclusions

In recent years, Venezuela has witnessed a rapid and massive exodus of over 2.3 million individuals who have migrated to Colombia, Chile, Argentina, Peru, and Ecuador. In Colombia, the influx of Venezuelan migrants has significantly increased since 2015, with an annual growth rate of 126%, compared to a 43% increase in 2014. The annual increments in migration rose to 174% and 188% in 2017 and 2018, respectively, consequently increasing the migrant population's presence initially in bordering departments. According to migration patterns, in the initial stage, migrants enter Colombia through regular points with the presence of Colombian Migration and irregular points along the 2219 km of shared land border. In the second stage, with more information available, the migratory flow shifts to urban centers with more substantial income-generating opportunities in Colombia. They continue to move south along the Ecuadorian border or north through the Darien Gap to enter Central America via Panama.

The unexpected surge in migratory flow into Colombia since 2016 led to regulatory changes, particularly the introduction of legal instruments to address the humanitarian emergency and subsequently facilitate the inclusion of Venezuelan-born migrants and returning nationals in social services, healthcare, education, and job placement. Despite regulatory changes, the migratory flow affected bordering departments with economic and social development gaps; many faced challenges in accessing resources and institutional constraints in responding to the restitution and effective enjoyment of rights for victims of the armed conflict in Colombia. The characteristics of the migration flow from Venezuela, primarily composed of working-age populations with education levels similar to the native population, pose challenges for accessing scarce resources in migration-affected areas, potentially resulting in competition for scarce job opportunities and xenophobia.

This paper investigates the consequences on the informal labor market segment in initially affected bordering departments due to the rapid and unforeseen arrival of migrants from Venezuela. Using different definitions of the informal workforce in a difference-in-differences model, the research tests the hypothesis that short-term labor migration increases (1) the number of employed individuals in companies with a workforce of fewer than five people, (2) the number of employed individuals not contributing to the social security system, and (3) the relative participation of informality in total employment. The main results indicate an expansion of the informal workforce segment, increasing the number of non-returned native workers in informality without significant increases in the participation of informality in total employment. The results remain robust across various samples in models adjusted for departmental, municipal, and individual-level data.

Several national and international stakeholders implementing income-generation alternatives in the border departments focusing on migrant employment services could find the research findings helpful in at least two aspects: first, fighting cultural stereotypes which result in native workers tending to see migration as a threat to their current job holdings requires evidence that shows that migrant workers contribute to economic growth and employment; second, promoting better public policies to take advantage of initial conditions that favor labor integration of migrant workers such as cultural and language similarities among natives and migrants works better when there is evidence of the migration's positive impacts. Although research findings deal with the informal segment of the labor market, usually depicted as a low labor productivity sector, having positive evidence of job creation for natives in the context of migration states the relevance of public policy design to articulate the migrants' qualifications with strategical economic sectors in Colombia. In other words, the lack of a relevant public policy design could end up increasing illegal markets of criminal hiring of migrants, increasing crime and cultural xenophobia against migrants.

There are four limitations to the findings' scope. First, having a pseudo panel of individuals instead of a longitudinal balanced panel of the same individuals becomes a limitation as long as the sample average of control variables could move along with the sample of individuals interviewed. However, descriptive analysis of the model variables shows a slight change in the control variables, proving grounds to consider that the sample design of the IHS survey fits the econometric model requirements. Second, the difference-in-differences (DID) model uses an unbalanced panel data method with fixed effects to control for characteristics that influence the migration effect but remain constant over time. In this regard, it only captures short-term effects that could vanish or prevail in the long term. Therefore, effects on consumption and investment caused by migration that might increase or decrease job opportunities for native workers are not in the research scope. However, the evidence of short-term effects could help inform public policy designs relevant to promoting migration's positive effects on economic welfare. Pathways for future research in this area include exploring distributional occupations' effects given demographic characteristics such as gender and age for providing evidence about changes in the distribution of occupations due to the migration effect. Moreover, the migration

effect on the labor market outcomes for native workers' cohorts could also provide a better picture of the migration's effect on different generations of native workers.

The third limitation refers to economic linkages among border cities in Colombia and Venezuela that might raise an argument about initial conditions regarding international trade factors in the econometric fixed factors model estimations, causing a sample bias. The selection of a control group does not adjust for such developmental variables because none of the departments and cities in the control group are close to Venezuela. However, modeling the developmental variable will require data aggregation for including import and export levels in volume and dollars as control variables in the cities and department models. Although we aggregate individual data to achieve metropolitan and departmental data, adding more control variables to the model will reduce the explanatory power because it reduces degrees of freedom. In this case, we decided to follow a parsimonious econometric identification principle, reducing explanatory variables to understand better labor market conditions rather than imposing more restrictions on the econometric tests. We acknowledge there is a potential bias that we are exploring in a different paper using a production function approach that could provide more data on the explanatory variables to examine the effects of developmental variables on the migration's effect on production and international trade among cities and departments in Colombia and Venezuela in the treatment group.

Finally, the fourth limitation concerns the influence of migrant networks on the migrants' supply shock exogenous assumption. Rapid and massive migration is random because there is no correlation between migrants' qualifications and job requirements in the labor markets in which they arrive. In this regard, information about job opportunities and job requirements that run through migrants' networks might have different quality levels influencing the homogeneity assumption, creating a correlation between the demographic characteristics of migrants and job openings. Therefore, different endogeneity forms could arise due to correlation patterns from the quality of information that runs through migrant networks. Appendix B acknowledges such possibilities of modeling the migrants' network effect when there are reliable data about it.

Author Contributions: Conceptualization, W.P.B.; methodology, W.P.B., C.D.C.R. and J.C.R.; software program, W.P.B. and J.M.D.; validation, W.P.B., J.C.R., C.D.C.R. and J.M.D.; formal analysis, W.P.B.; investigation, W.P.B.; resources, J.M.D.; data curation, J.M.D.; writing—original draft preparation, W.P.B., C.D.C.R. and J.C.R.; writing—review and editing, W.P.B.; visualization, W.P.B.; supervision, W.P.B.; project administration, W.P.B.; funding acquisition, W.P.B. All authors have read and agreed to the published version of the manuscript.

Appendix A

Appendix A presents the selection criteria for the control group. The departments bordering Venezuela are La Guajira, Cesar, Norte de Santander, Boyacá, Arauca, and Vichada. Of these departments, La Guajira, Cesar, Norte de Santander, and Boyacá are part of the household sample surveyed by the IHS implemented by NADS in Colombia. For this reason, the departments of Arauca and Vichada are not included in the treatment group, even though, according to available diagnostics on irregular migration, particularly in the case of the department of Arauca, they receive a significant influx of irregular migrants from Venezuela.

The criteria selected for establishing the treatment group were as follows:

- Non-bordering department.
- Department with a development environment measured by DNP similar to the development environment of the departments in the treatment group.
- Department has a lower incidence of immigration from Venezuela.

Table A1 describes the control group and treatment group according to the established criteria.

Table A1. Control group and treatment group.

Group	Departments	Average Rating Development Environment	Immigration Incidence 2013–2018
Treatment	Boyacá	0.47	8405
	Cesar	0.41	71745
	Norte de Santander	0.39	115447
	La Guajira	0.27	148265
	Group Average	0.38	
	Total Group Migration		343863
Control	Caldas	0.42	6434
	Choco	0.28	765
	Huila	0.44	5072
	Tolima	0.39	7662
	Group Average	0.38	
	Total Group Migration		19933

Source: authors' estimates based on the Typology of Departmental Development Environments of the DNP and the Migration Module of the GEIH.

The average score for the development environment in the treatment group is 0.38, similar to the average score in the control group. The Development Environment Index is elaborated by the DNP for departments and municipalities, measuring the economic development capacity across six dimensions: (1) urban dimension, (2) economic dimension, (3) quality of life dimension, (4) environmental dimension, (5) security dimension, and (6) institutional dimension. A detailed description of the index and the variables considered in each dimension can be found in Carmona et al. (2015). The highest index within the treatment group is in Boyacá (0.47), while the lowest is in La Guajira (0.27). On the other hand, the highest index value in the control group is in Huila (0.44), and the lowest is in Chocó (0.28).

The departments of Caldas, Chocó, Huila, and Tolima, which have average development environment indicators similar to those of the departments of La Guajira, Cesar, Norte de Santander, and Boyacá in the treatment group, do not share a border with Venezuela and, therefore, do not experience the same level of migration flow in the initial phase as that observed in the bordering departments. The total migration between 2013 and 2018 for the departments in the treatment group amounted to 148,265 people, while the influx of migration in the departments in the control group reached 19,933 people during the same period.

Appendix B

Appendix B describes the potential bias coming from the existence of migrant networks in which there is information about the native labor market's job openings and job requirements. The research assumes there is no correlation between migrant skills and work experience with job openings in the native labor market because of data constraints in measuring the existence of migrant networks and the information quality that runs through the migrant networks. In this regard, hidden forms of correlation might cause a bias on the migrants' effect on the native informal workforce. Figure A1 shows how the research acknowledges the potential bias that could arise due to a migrants' network effect.
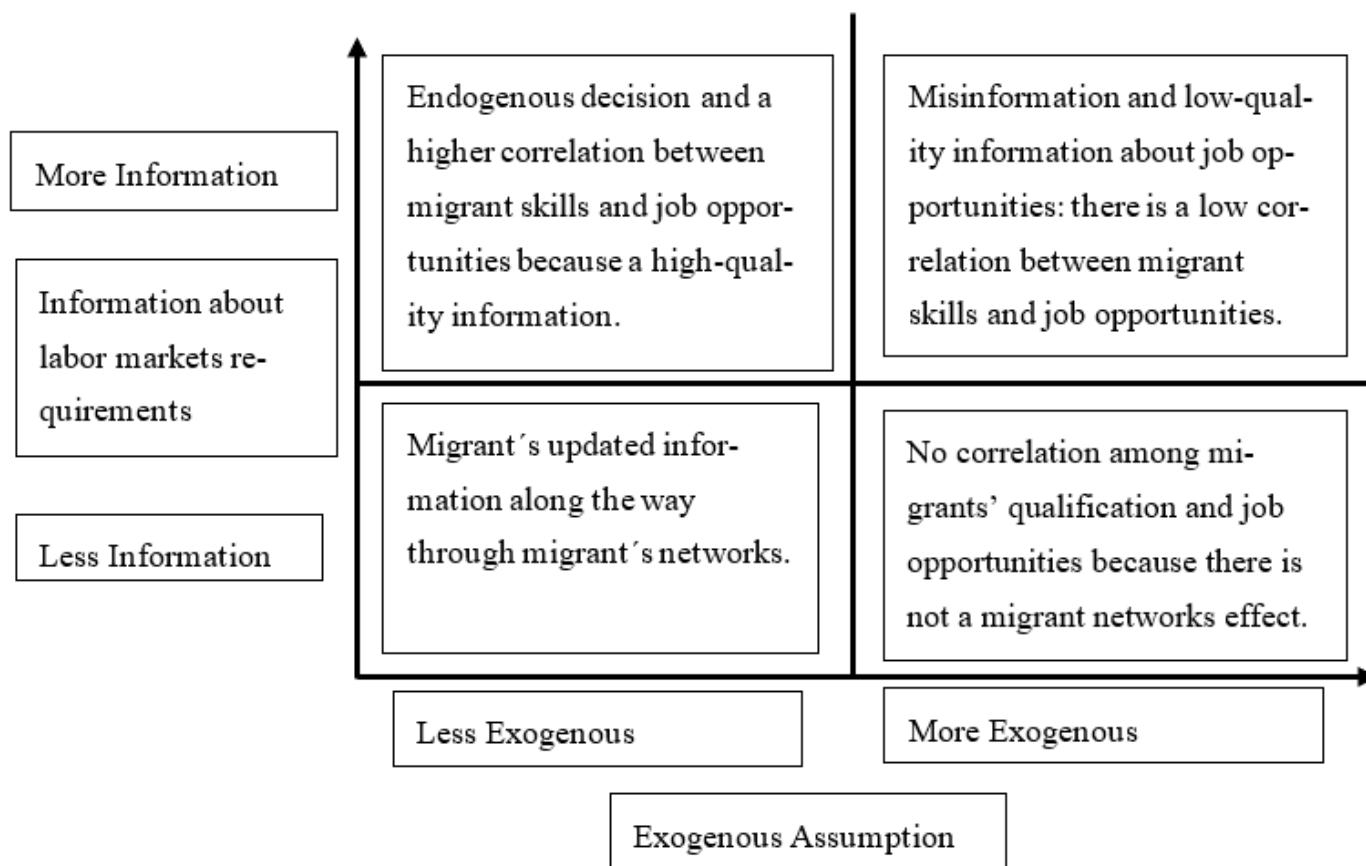
	Endogenous decision and a higher correlation between migrant skills and job opportunities because a high-quality information.	Misinformation and low-quality information about job opportunities: there is a low correlation between migrant skills and job opportunities.
	Migrant´s updated information along the way through migrant´s networks.	No correlation among migrants' qualification and job opportunities because there is not a migrant networks effect.

More Information

Information about labor markets requirements

Less Information

Less Exogenous

More Exogenous

Exogenous Assumption

Figure A1. Migrant networks and endogenous information about job openings.

References

Altonji, Joseph G., and David Card. 2019. The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives. In *Interdisciplinary Perspectives on the New Immigration*. Edited by M. S.-H. Suarez. London: Routledge, p. 128. [CrossRef]

Aydemir, Abdurrahman, and George J. Borjas. 2007. Cross-country variation in the impact of international migration: Canada, Mexico, and the United States. *Journal of the European Economic Association* 5: 663–708. [CrossRef]

Bağır, Yusuf Kenan. 2018. Impact of the Syrian refugee influx on Turkish native workers: An ethnic enclave approach. *Central Bank Review* 18: 129–47. [CrossRef]

Balkan, Binnur, and Semin Tumen. 2016. Immigration and prices: Quasi-experimental evidence from Syrian refugees in Turkey. *SSRN Electronic Journal*. [CrossRef]

Becker, Sascha O., and Andreas Ferrara. 2019. Consequences of forced migration: A survey of recent findings. *Labour Economics* 59: 1–16. [CrossRef]

Bernal, Raquel B., and Ximena Peña. 2011. *Practical Guide for Impact Evaluation*, 1st ed. Bogotá: Universidad de los Andes, p. 235.

Borjas, George J. 2003. The labor demand curve is downward sloping. *The Quarterly Journal of Economics* 118: 1335–74. [CrossRef]

Borjas, George J. 2014. *Immigration Economics*. Cambridge: Harvard University Press.

Callaway, Brantly, Tong Li, and Tatsushi Oka. 2018. Quantile treatment effects in difference in differences models under dependence restrictions and with only two time periods. *Journal of Econometrics* 206: 395–413. [CrossRef]

Cameron, Adrian Colin, and Pravin K. Trivedi. 2009. *Microeconometrics Using Stata*. College Station: Stata Corp LP, 677p.

Card, David. 1990. The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial and Labor Relations* 43: 245. [CrossRef]

Card, David. 2001. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics* 19: 22–64. [CrossRef]

Card, David, and Giovanni Peri. 2016. Immigration economics by George J. Borjas: A review essay. *Journal of Economic Literature* 54: 1333–49. [CrossRef]

Carmona, C., D. Supelano, and I. Osejo. 2015. Tipologías Departamentales y Municipales: Una Propuesta para Entender las Entidades Territoriales Colombianas. Colaboración con el DPS. Available online: https://colaboracion.dnp.gov.co/CDT/ (accessed on 5 October 2023).

Caruso, German, Christian Gomez Canon, and Valerie Mueller. 2019. Spillover effects of the Venezuelan crisis: Migration impacts in Colombia. *Oxford Economic Papers* 73: 771–95. [CrossRef]

Castillo, Robayo, William Prieto, and Jacobo Campo. 2022. *Impact of Migration on Employment and Wage Distribution in Colombia*. Working Paper. Bogotá: Faculty of Economic and Administrative Sciences, Catholic University of Colombia.

Ceritoglu, Evren, H. Burcu Gurcihan Yunculer, Huzeyfe Torun, and Semih Tumen. 2017. The impact of Syrian refugees on natives' labor market outcomes in Turkey: Evidence from a quasi-experimental design. *IZA Journal of Labor Policy* 6: 5. [CrossRef]

Esen, Oğuz, and Ayla Oğuş Binatli. 2017. The impact of syrian refugees on the turkish economy: Regional labour market effects. *Social Sciences* 6: 129. [CrossRef]

Fallah, Belal, Caroline Krafft, and Jackline Wahba. 2019. The impact of refugees on employment and wages in Jordan. *Journal of Development Economics* 139: 203–16. [CrossRef]

Fan, Yanqin, and Zhengfei Yu. 2012. Partial identification of distributional and quantile treatment effects in difference-in-differences models. *Economics Letters* 115: 511–15. [CrossRef]

Friedberg, Rachel M. 2001. The impact of mass migration on the israeli labor market. *The Quarterly Journal of Economics* 116: 1373–408. [CrossRef]

Llull, Joan. 2018. The Effect of Immigration on Wages: Exploiting exogenous variation at the national level. *The Journal of Human Resources* 53: 608–62. [CrossRef]

Mora, Jhon James, Andres Cuadros-Menaca, and Juan Tomas Sayago. 2022. South-South migration: The impact of the Venezuelan diaspora on Colombian natives' wages and hours worked. *Applied Economics Letters* 30: 884–91. [CrossRef]

Morales, Fernando, and Martha Denisse Pierola. 2020. *Venezuelan Migration in Peru: Short-term Adjustments in the Labor Market*. IDB Working Paper Series No. IDB-WP-1146. Washington, DC: Inter-American Development Bank. [CrossRef]

Olivieri, Sergio, Francesc Ortega, Ana Rivadeneira, and Eliana Carranza. 2022. The Labour Market Effects of Venezuelan Migration in Ecuador. *The Journal of Development Studies* 58: 713–29. [CrossRef]

Ottaviano, Gianmarco I. P., and Giovanni Peri. 2021. Rethinking the Effect of Immigration on Wages. *Journal of the European Economic Association* 10: 152–97. [CrossRef]

Pedrazzi, Julian, and Leonardo Peñaloza-Pacheco. 2021. *Heterogeneous Effects of Forced Migration on Female Labor Supply*. CEDLAS Working Papers 0274. La Plata: Universidad Nacional de La Plata.

Peñaloza-Pacheco, Leonardo. 2022. Living with the neighbors: The effect of Venezuelan forced migration on the labor market in Colombia. *Journal for Labour Market Research* 56: 14. [CrossRef]

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb. 2019. Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal* 19: 4–60. [CrossRef]

Rosales, R., J. Perdomo, C. Morales, and J. Urrego. 2013. *Fundamentos de Econometría Intermedia: Teoría y Aplicaciones*. Bogota: Universidad de los Andes, 405p.

Santamaria, Julieth. 2020. *When a Stranger Shall Sojourn with Thee: The Impact of the Venezuelan Exodus on Colombian Labor Markets*. Working Paper 51422. Minneapolis: Department of Applied Economics, University of Minnesota.

Tumen, Semih. 2015. *The Use of Natural Experiments in Migration Research*. Bonn: IZA World of Labor, p. 191. [CrossRef]

Tumen, Semih. 2016. The Economic Impact of Syrian Refugees on Host Countries: Quasi-Experimental Evidence from Turkey. *The American Economic Review* 106: 456–60. [CrossRef]

World Bank. 2018. Migration from Venezuela to Colombia: Impacts and Short- to Medium-Term Response Strategy. Available online: https://reliefweb.int/sites/reliefweb.int/files/resources/131472SP.pdf (accessed on 5 October 2023).