Trabajo presentado como requisito para optar por el título de Magister en Economía

Título tesis:

The effect of Venezuelan migration on city-level productivity in Colombia

Autor:
Mario Alejandro Nieves Buitrago

Director, Tutor:
Juan M. Gallego

Universidad del Rosario

Facultad de Economía
Maestría en Economía
Universidad del Rosario
Bogotá, Colombia
2022

The effect of Venezuelan migration on city-level productivity in Colombia
El efecto de la migración venezolana en la productividad municipal en Colombia

Mario Alejandro Nieves Buitrago†
Thesis

Abstract

This paper evaluates the effect of Venezuelan migration on productivity at the city level (municipality level) in Colombia between 2013 and 2018. This relationship faces endogeneity on two fronts. On the one hand, productivity estimation faces an omitted variables bias, while the relationship between migration and productivity may have dual causality. These problems are solved using a control function, and instrumental variables approach (Bartik instrument), respectively. The results indicate that productivity at the municipality level has not been impacted by Venezuelan migration. Beyond productivity, there are no effects on urbanization and agglomeration effects. These results can be explained by the fact that migrants face hiring barriers and must resort to informal jobs, the mismatch between the growing labor force and labor demand, and the misallocation of production factors. However, there are positive and significant effects on the specialization of intermediate cities.

Resumen

Este trabajo evalúa el efecto de la migración venezolana en la productividad de las ciudades (municipios) de Colombia entre 2013 y 2018. Esta relación se enfrenta un problema de endogeneidad en dos frentes. Por un lado, la estimación de la productividad padece de variables omitidas, mientras que la relación entre la migración y la productividad puede tener doble causalidad. Estos problemas se resuelven empleando un enfoque de función de control y de variables instrumentales (instrumento Bartik), respectivamente. Los resultados indican que la productividad a nivel de municipio no ha sido impactada por la migración venezolana. Más allá de la productividad, no hay efectos en los efectos de urbanización y aglomeración de la ciudad. Estos resultados pueden explicarse por el hecho de que los migrantes enfrentan barreras de contratación y deben recurrir a trabajos informales, por la no correspondencia entre la creciente fuerza laboral y la demanda de trabajo y finalmente la *misallocation* de factores producción. Sin embargo, sí hay efectos positivos y significativos en la especialización de las ciudades intermedias.

Keywords— Migration, Productivity, Instrumental Variables, Control Function, Cities
JEL F22 O47 R12

I am deeply grateful to my advisor Juan Miguel Gallego for his guidance through this process. I am also grateful to Iván de las Heras, Mateo Cardona, and Daniel Prieto for their thoroughly reading and deepful comments in this work.

†Universidad del Rosario. email: mario.nieves@urosario.edu.co

1 Introduction

In the face of war, famine, and scarcity, human beings have resorted to migration to safeguard their income prospects and opportunities, both individually and for their offspring. Migrants, like the receiving countries, benefit from this exchange: the former find a place to settle, build their life projects and offer their labor skills, while the latter offer development opportunities and receive human capital without assuming the cost of their training. According to economic theory, the impact of migration on economic growth is positive, so removing migration barriers brings benefits to receiving countries in the long run (Portes, 2019).

Immigrant workers can fill gaps where skills are in short supply and bring ideas, such as knowledge of foreign markets, new ways to do things, and new things to do (Fabling, Maré, & Stevens, 2022). These benefits may be significant for a country with little migration compared to other Latin American countries like Colombia. Indeed, Colombia's internal dynamics have led to a negative net migration rate. Indeed, approximately 4.7 million Colombians reside outside the country during the last census (2018). Historically, this was the second-largest diaspora in Latin America until Venezuela took over this position (Vásquez-DeKartzow, Castillo, & Durán, 2015).

As per local labor surveys, it is estimated that there are more than 1.9 million Venezuelans residing in Colombia. The massive influx of Venezuelans is currently among the most widely discussed topics in the country in terms of public policy and presents significant challenges in labor, public spending, and quality of life. Moreover, even though Colombia has never received a migratory wave of this magnitude, the few historic foreign enclaves have greatly influenced the receiving cities and their business development (Cárdenas & Mejía, 2006), as the case of the migration of Lebanese and Germans in Barranquilla and Japanese to Valle del Cauca.

Although many studies have recently studied the effect of migration, much of the focus is on the impact of migrant labor on native workers' wages and employment, and there is no extensive literature on the relationship between migration and productivity. In this topic, Colombia has a stagnant trajectory. The lag in Total Factor Productivity (TFP) has subtracted 0.4% of growth from the country over the century, as estimated by the Private Competitiveness Council (CPC). If Colombia's TFP had Peru's behavior, GDP per capita would have been 22% higher in 2019. The development of countries with impressive economic behaviors, such as South Korea or Singapore, coincided with TFP growth (Nguyen, 2018). In this sense, looking for ways to enhance productivity should be a priority in development policy.

Theoretically, TFP is the residual of output that is not explained in the amount of capital and labor in the production process, and given that labor is one of two main production factors, it is worth asking whether Venezuelan migration has allowed Colombia to obtain the fruits of immigration that have so far eluded, both in terms of migration and productivity. These two variables are connected in many ways. First of all, as Fabling et al. (2022) point out, the idea that immigration increases only labor supply, not labor demand, does not appear to be supported by evidence. Secondly, the capital stock is not fixed and will shift in response to changes in the level and composition of labor supply (Ottaviano & Peri, 2012). In addition, net immigration increases the total labor supply, increasing the return to capital and increasing capital investment, productivity, and labor demand (Friedberg & Hunt, 1995). This impact will be even higher if migrant workers are complementary with capital and lower if they are substitutes (Fabling et al., 2022).

Venezuelan-born people are currently a vital component of the Colombian population as they represent approximately 3.4% of the total population. Nonetheless, any survey in Colombia collects information about foreign workers employed in the firms. For this reason, this research consists of two stages: i) measuring firm productivity and ii) evaluating migration's impact on city-level productivity (for this purpose, city and municipality are interchangeable terms in Colombia). The common channel between them is city localization, as foreign migration can be observed at the city level using the last censuses (2005 and 2018), and productivity can be measured as an aggregate of firm-level productivities. According to Krugman and Obstfeld (2009), firms' productivity depends not only on internal characteristics but also on external forces, such as the concentration of sectors in a region that may lead to positive externalities, specialized suppliers, labor market pooling, and knowledge spillovers. Therefore, counting for city characteristics and foreign migration makes a comprehensive set of productivity boosters.

In the firm stage, the Production Function is estimated through a control function to avoid endogeneity caused by non-observed productivity. In order to connect firms and municipalities stages, TFP for firms in the municipality j will be added to obtain aggregated productivity at the municipality level. In the second stage, aggregated municipality-level productivity will be decomposed into city-level characteristics, including migration and relevant city features identified in the literature (Glaeser, 2011), such as crime, size, and territorial management. In the second stage, an endogeneity problem between migration and productivity arises, given that migrants self-select into the cities that offer them the best economic opportunities. Therefore, this research proposes the *Bartik* shift-share instrument to circumvent this double-causality problem. The empirical strategy exploits this municipal variation in intensity, given that there are municipalities with previously established Venezuelan networks prior to the migration crisis compared to other municipalities that received little relative migration (Rozo & Vargas, 2021).

This study aims to measure the effect of migration on city-level productivity. The results of this research indicate that city-level productivity has not been positively affected by Venezuelan migration. These results can be explained by the fact that migrants face hiring barriers and must resort to informal jobs and the mismatch between the growing labor force and labor demand. There is no effect on value-added and total product. However, there is a positive effect on city specialization, consistent with the perspective of Fabling et al. (2022). The rest of the paper is structured as follows. Section 2 presents the literature review. Section 3 gives an overview of the data coming from firms, municipalities, and migrants. Section 4 presents the empirical strategy. Section 5 reports the Statistic and main challenges and findings regarding the data. Section 6 presents the results, including the robustness analysis and other considered spillovers of migration. Finally, the last section discusses and concludes.

2 Literature review

The topic of migration has received considerable critical attention in developed countries. However, the literature does not have a unanimous perspective on the effect of migration on labor or productivity. Borjas, Freeman, and Katz (1996) argue in early studies that there are two opposing views about how immigrants affect the labor market opportunities of natives in the U.S. On the one hand, migration has a negative effect based on the competition for the same jobs. On the other one, migrants' and natives' services complement each other. The authors also assert that the first one is more likely correct. There is also a range of potential within-firm complementarities to consider (Fabling et al., 2022). As the authors point out, migrants may be complementary with (or substitutes for) other factors of production.

From a positive perspective, Jahn and Steinhardt (2018) argue that to the extent that immigration has a different composition than the native population, this will change the skill composition of the labor pool as a whole. Excess labor of lower-skilled local workers can be absorbed into the local market by expanding production. In the same vein, according to Tribin (2020), migrants from Venezuela are substantially younger and have similar levels of education compared to the non-migrant Colombian population. Further, working-age immigrants increase the tax base and reduce per capita medical and retirement costs, at least in short to medium-term (Fabling et al., 2022).

Orrenius and Zavodny (2012) assert that migration is an opportunity for receiving countries, as migrants receive wages below the legal minimum, which benefits employers and increases the number of hired workers. In the same sense, Hanson (2010) argues that immigration increases productivity when migrants have high levels of education. In the U.S. case, the author finds that foreign engineers create new production technologies that make capital and labor more productive. Second, labor-intensive industries increase profits, hire more workers, and use more land and capital. The production also expands in labor-intensive sectors and necessarily declines in others.

Notably, changes in competitiveness (and productivity) are sector-specific, generating differential income changes according to the type of factor.

Martínez and Muñoz (2020) argue that migration can propel growth as its main barriers are high labor, transportation, and energy costs, and migration flows can reduce labor costs. In addition, considering the composition of Venezuelan migration and the fact that the arrival of low-skilled workers has increased during the last few years, migration may lead to expanding the production (provision) of certain products (services) that intensively use low-skilled labor (Borjas, 1995). In this transmission channel, the sector's expansion may increase demand and increase wages. In addition, gains from migration will be positive for the world economy since moving labor from low-productivity to high-productivity countries improves allocative efficiency (Hanson, 2010).

There could be adverse effects from migration. An increase in unskilled migrants may disincentivize investment in labour-replacing capital (e.g., reduce uptake of automation) (Fabling et al., 2022). In this case, immigration of labor at different skill levels may cause firms to invest and (or) disinvest in the capital of different types (Edo, 2019; Lewis, 2011, 2013; Longhi et al., 2010, as cited in Fabling et al. (2022)). In addition, as Borjas (1995) argues, the simple static effect of an increase in labor supply leads to a decline in wages, all other things equal. However, according to Fabling et al. (2022), all other things are seldom equal because migrants demand food, clothing, housing, and other amenities. These products and services must be produced, often domestically, increasing the local demand for labor.

Regarding city productivity, Borjas et al. (1996) sustain that immigration generates effects nationwide instead of only in those cities with more pronounced migration since labor and capital are mobiles between cities. The authors also argue that native workers and immigrants would leave a city and work elsewhere if economic conditions deteriorated and that labor migration would trigger capital to leave with them. Therefore, this research will use a city-level model to explore aggregated effects on productivity. In addition, there is also a broader composition effect at the national level (Fabling et al., 2022). Immigration can raise the average productivity of the whole economy because migrants are typically more likely to be of working age than the native-born population (Jaumotte, Koloskova, & Saxena, 2016).

Productivity is not only a product of firms' features, as firms' location also plays an essential role (Krugman & Obstfeld, 2009). The authors assert that firms tend to be located within a short distance of each other to reap the benefits of external economies (e.g., California's Silicon Valley and New York financial sectors). In addition, Casas and Balat (2018) highlight that standard production function estimates omit city-related determinants since productivity is estimated at the firm level. The authors argue that agglomeration economies emphasize external forces to the firm. Therefore, it is crucial to reveal systematic relationships between firm-level productivity and the cities' features using firm and city data.

Recent studies have analyzed the impact of Venezuelan migration on different spheres of the Colombian economy, mainly on the labor market. On the one hand, Santamaría (2019) finds insignificant reductions in informal sector wages due to Venezuelan migration. In contrast, Caruso, Gómez, and Mueller (2020) determines a 10 percent reduction in wages in this segment as migration increases by one percent. Along the same lines, Pulido and Varon (2020) observe that migration generates misallocation in the short term. In addition, reducing barriers to migrants would raise labor productivity by 0.9%.

In the labor market, Bonilla, Morales, Hermida, and Flórez (2020) find that the migratory flow increases unemployment among immigrants but does not significantly affect natives. In this group, job losses are mainly concentrated in the self-employed segment and generally are more pronounced for women, young people, and individuals with low labor qualifications. On the other hand, different studies have incorporated the endogeneity of migration in different dependent variables, such as electoral participation and political orientation. Rozo and Vargas (2021) find that the increase in Venezuelan migration increases electoral participation and votes for right-wing parties.

The authors use the enclave instrumental variables approach since this method is useful to circumvent the double causality between migrants' location choices and the dependent outcomes.

Recently, researchers have started their scrutiny of the effect of migration on productivity in Colombia. For instance, Martínez and Muñoz (2020) using Gross Domestic Products data of 24 departments from 2014 to 2018 found that Venezuelan immigration has a significant negative effect on departmental productivity in Colombia, reducing its growth by 48.46%, as it increases the proportion of hours worked by less-educated workers, and highly educated workers become less productive than less-educated workers. This process occurs through a *skill bias* in which the share of hours worked by less educated workers increases, and highly educated workers become less productive than less educated workers.

This paper contributes to the literature by providing input to the growing body of research on the effects of Venezuelan migration on different dimensions of the Colombian economy. Likewise, studies for south-south migration flows are underrepresented in the literature. Therefore, exploring the main transmission channels in developed countries may present a novel point of view. There are other significant differences in Colombian-Venezuelan migration flows to canonical migration studies, such as cross-border relations that produce relatively low cost of migration, similar cultures, and a receiving country with high levels of informality. In addition, the change of nomenclature in financial statements to the Superintendence of Corporations *Supersociedades* represents significance in constructing production factor series. In this sense, this is the first article that uses both surveys, as it was studied under the previous accounting standards in (Casas & Balat, 2018).

3 Data

In order to answer this paper's central question, it is necessary to collect data on the following three fronts: firms, municipalities, and individuals (including migrants). The databases used on each front will be presented below. The study period is 2013-2018 since it is the year the Venezuelan migration crisis intensified.

3.1 Firms

First, to identify TFP using firm-level data, this investigation employs two information-rich data panels to construct the production function. These databases are the Enterprise Risk Information System (*SIREM*, in Spanish) and the Integrated Corporate Information System (*SIIS*, in Spanish). These surveys are collected by *Supersociedades* and contain the financial statements at different levels for the companies supervised by this entity. A strength of this source is that it has a municipal identifier, unlike the Annual Manufacturing Survey (*EAM*). As we observed in (as is observed in Figure 1), the bulk of Colombia's economic activity is concentrated in the highlighted municipalities, namely, the main metropolitan areas.

This source provides disaggregated information for more than 154 municipalities. SIREM was implemented from 1999 to 2015, whereas SIIS was implemented from 2016 onwards and is the updated version of SIREM. The reason for this update is the nomenclature change of accounting reports. The Single Chart of Accounts (Plan Único de Cuentas in Spanish, PUC) was presented until 2015. From 2016, the International Financial Reporting Standards (NIIF in Spanish) came into force. This work also collects a matching exercise of the variables necessary to construct the production function and Total Factor Productivity using each nomenclature.

Table (1) shows which items were used to construct the production function. It is worth mentioning that there is no correlative table between the Single Chart of Accounts (PUC) and the International Financial Reporting Standards (NIIF). In this paper, the operating income from the *SIREM* survey was used for *production*, which in *SIIS* is called income from ordinary activities, which is its analog. On the other hand, the items property, plant and equipment, and labor liabilities have the same naming in both surveys and were used for the *capital* and *labor* variables, respectively. Finally, the *materials* factor is necessary for estimating productivity using the control function method for recovering unobserved productivity. This variable was constructed as the sum of inventories in *SIREM* called *short* and *long-term inventories*. In *SIIS*, they are named *current* and *non-current inventories*, with current inventories being short-term and non-current inventories being long-term. More information regarding variables is contained in Appendix - Table (10).

Table 1: Items used to construct the Production Function.

	SIREM	**SIIS**
Product	Operational revenue	Revenue from ordinary activities
Capital	Property, plant and equipment	Property, plant and equipment
Labor	Labor liabilities	Labor liabilities
Materials	Short-term + Long-term inventories	Current + Non-current inventories

3.2 Migrants

Regarding migrants, the Great Integrated Household Survey (GEIH in Spanish) will be used to characterize the population. It is the primary reference on Colombia's labor market, it includes a module on migration since 2013. This module asks where the respondent was born and where he/she lived one and five years ago. This variable allows us to construct the series of Venezuelan migrants in the study period using expansion factors. One of the strengths of using the GEIH is that it does not present self-selection in the registration process. The number of migrants observed each year presents fewer jumps concerning other sources such as the Administrative Registry of Venezuelan Migrants (RAMV) and Migración Colombia, as pointed out by Tribin (2020).

This paper will also employ the National Population and Housing Census (2005 and 2018 waves). The shift-share instrument constructed to solve the endogeneity problem between productivity and migration considers the proportion of Venezuelan-born for each municipality in Colombia in 2005. In order to evaluate the relevance of the instrument, it will be used as a regressor variable of the actual number of foreign-born people reported in the 2018 Census. In addition, there is no source of information containing the number of Venezuelans per municipality for the whole period (2013-2018). Thus, this exercise proposes a valid approximation.

3.3 Municipalities

To characterize the municipalities, the Municipal Database of the *Universidad de los Andes'* CEDE provides exhaustive information on the municipalities between 1993 and 2018, accounting for the control variables included in the city estimation. These variables are collected in six different modules: i) Good governance; ii) Agriculture and land; iii) General characteristics; iv) Conflict and violence; v) Education, and vi) Health and services.

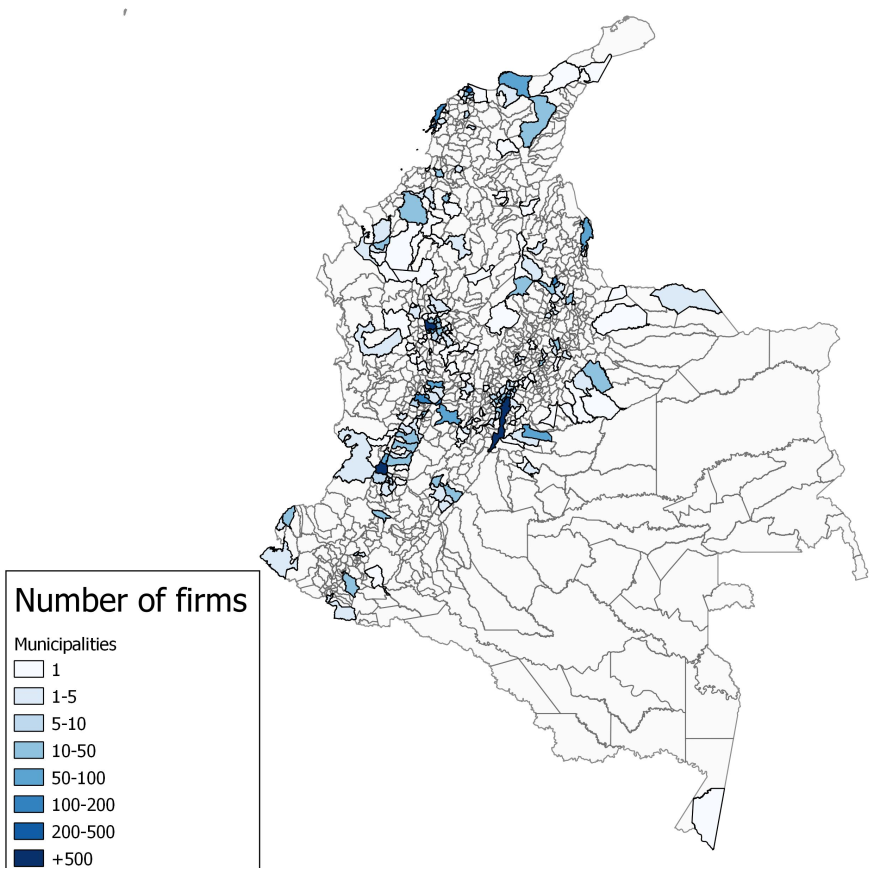


Figure 1: Firms count by city in 2018.

4 Empirical strategy

4.1 Firms

The effect of migration on business productivity will be estimated in two stages: (i) firm productivity and (ii) municipalities. The first stage uses the Production Function to estimate Total Factor Productivity (TFP) at the firm level. TFP is the residual of output that is not explained in the amount of capital and labor in the production process, and this residual is a black box representing technical change that leads to sustainable development (Nguyen, 2018). Its use in the empirical analysis plays a key role not only in the dispersion of productivity but also in other areas such as technological change based on human capital, estimation of economies of scale, evaluation of the effect of new technologies, learning by doing, among others (Aguirregabiria, 2018). The dependent variable at the second stage at the city level is city-level productivity and employs a migration instrument to avoid endogeneity along with other city characteristics.

Aguirregabiria (2018) argues that the estimation of the production function (PF) faces a simultaneity problem, where unobserved productivity ω_{it} is known to the firm when it decides the inputs it will use in its production, and these inputs are correlated with ω_{it}. Therefore, using a Cobb-Douglas specification would yield biased estimates. Hence, Olley and Pakes (1996) propose a Control Function methodology for estimating TFP. Unobserved productivity can be recovered as the inverse of the demand for flexible inputs (labor or intermediate materials) by semiparametric estimates referring to the production function. TFP can then be calculated as the difference between the observed output minus the output adjusted to the estimate, as presented in Equation (1).

$$\hat{\omega}_{it} = y_{it} - \hat{\beta}_l l_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_m m_{it} \tag{1}$$

Where y_{it} corresponds to the observed product whereas $\hat{y}_{it}$ is the estimated product. The following three variables (l_{it}, k_{it} and m_{it}) are wages, capital and intermediate materials, respectively. The term $\hat{\omega}_{it}$ obtained in 1 is defined as productivity estimated at the firm level. Therefore, to move to the second stage (municipalities), the firms' productivity will be aggregated and weighted by each firm's output (s) to form the municipal indicator ($\hat{W}_{mt}$) as shown in Equation (2).

$$\hat{W}_{mt} = \sum sy_{imt}\hat{\omega}_{it} \tag{2}$$

In order to clean the data and present relevant indexes, many procedures were carried out. First of all, there were specific deflators for each production's factor. These indexes were obtained from the Domestic Supply Price Index and were related as follows: Final consumption was used to deflate operating income (y); Intermediate consumption for materials (m), and Capital goods for the capital (k). The Consumer Price Index (CPI) was used to deflate the labor factor (l). On the other hand, the CAPEX methodology was used to obtain the gross fixed capital formation, which allows measuring the investment in capital or fixed assets made by a company to acquire, maintain or improve its non-current assets using the financial statements provided by *Supersociedades*. In addition, the detail of control function is explained on the Appendix (7.4).

4.2 Municipalities

In this stage, the objective is to explain the variability of productivity at the municipality level into different characteristics, including migration. Therefore, the term $\hat{W}_{mt}$ from Equation (2) corresponds to the dependent variable of interest in this research (city-level productivity). The functional form of this variable is explained in Equation (3).

$$\hat{\mathbf{W}}_{mt} = BX_{mt} + \Phi Migration_{mt} + \mu_m + \mu_t + \epsilon_{mt} \tag{3}$$

First, indexes m and t indicate that panel data is indexed at the municipal level, and its time-frequency is annual. Where X_{mt} corresponds to city characteristics collected from the CEDE municipal panel, $Migration_{mt}$ to migration influx. It is precisely the associated parameter Φ that measures the effect of the research question. In addition, to include unobserved attributes that do not vary over time and include time-fixed effects common to all municipalities, fixed effects for municipality and year are included in μ_m and μ_t, respectively. Estimate Equation (3) without instrumenting migration has a considerable drawback: there may be double causality, and migration may be affected by productivity since migrants will self-select in municipalities where they perceive better income prospects. For this reason, a shift-share instrument is employed.

In addition to the temporal variation of Venezuelan migration over time, it is necessary to consider its geographic variation. Venezuelan migration was received at a national scale, and some municipalities were more exposed to the migratory flow. In this sense, different instruments resolve the endogeneity existing between migration and productivity. First, the past localization of migrants can be used as an instrument since motives to migrate at time $t-k$ are correlated with migration at time t but are not probably correlated with productivity at t.

This approach is valid since migration prior to Venezuela's crisis was based mainly on family and neighbor relations rather than escaping a social and economic crisis. Using the number of Venezuelans projection for each year will allow a relevant instrument to the endogenous variable (present migration) but not the dependent variable (productivity, as it is exogenous). It is worth asking for an exclusion restriction. As per Daly (2021), Venezuelan migrants arrive because of economic conditions that also affect productivity. However, most Venezuelans that currently migrate to Colombia do it because it is the only way out to leave Venezuela and Colombia is the closest country, so they do not arrive mainly following the dynamics of the country.

The control variables were selected according to the literature review to establish whether there were any other boosting or mitigating factors. Glaeser (2011) argues that the productivity of cities varies according to crime indicators. For this reason, crime rates per 100,000 inhabitants will be used for homicides and thefts. The author also argues that management indicators affect productivity, so the municipality's tax revenues (per capita) and income from the General System of Participation (SGP, per capita) are used. Finally, exploiting the fact that both SIREM and SIIS include a municipal indicator, the distance to the departmental capital is also included, to the extent that productivity dynamics are not the same in the departmental capital as in its surrounding metropolitan area.

4.3 Measuring migration

Venezuelan migration has varied exponentially over time, and obtaining the closest estimation to reality is a pivotal task. Several studies have sought to determine the best approximation, and the migration series is constructed as follows in this paper. First of all, the 2005 National Population and Housing Census contains a question about the country of origin of the respondents, where only those coming from Venezuela (identified with code 850 in the country codification of the National Tax and Customs Directorate - DIAN) are considered, with this question in mind, the number of Venezuelan residents in each municipality -and therefore, in Colombia- is measured.

Subsequently, the fifth module of the Large Integrated Household Survey (GEIH) is employed. This survey contains a question about which country the respondent lived in a year ago, five years ago, and where he/she was born. This last question is selected employing the expansion factors, and with all the year surveys, the estimated number of Venezuelans residing in Colombia is obtained. This source of information has a strength compared to the official sources (Administrative Registry of Venezuelan Migrants and Migration Colombia), which have a drawback of self-selection (Tribin, 2020) and more abrupt jumps in the presentation of the series.

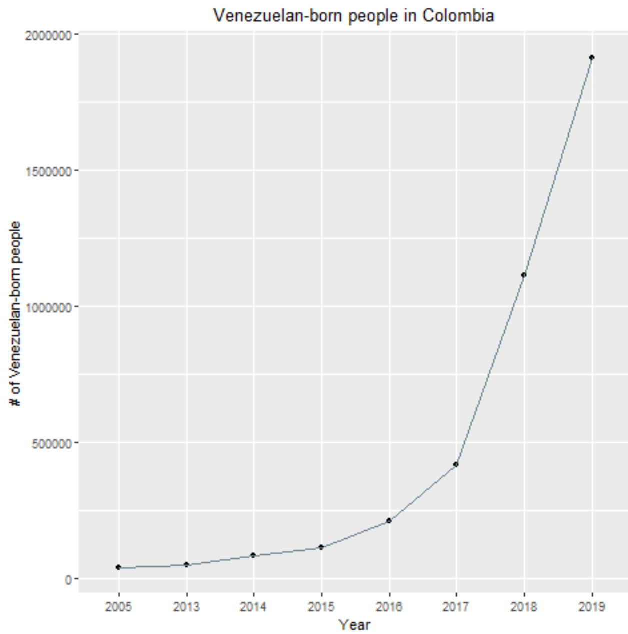


Figure 2: Venezuelan-born people in Colombia. Source: GEIH.

Figure (2) shows the number of Venezuelan residents in Colombia during the study period, and Table (2) contains the data depicted on the graph. We observe that the variation from 2013 to 2014 is even bigger than from 2005 to 2013. This colossal variation shows the magnitude of migration that would increase exponentially over the next five years. As the scope of this research is both municipal migration and municipal productivity, the variations in column (3) will be used to project the base population (the year 2005) up to the study period.

In Figure (3), we observe the composition according to the city of residence of Venezuelans in the fifteen main destinations. We observe that in 2005, neighborhood relations predominate as the main motive of residence since Cúcuta occupies the second position and its neighboring municipalities (Villa del Rosario, Los Patios) are all placed

Table 2: People born in Venezuela. Source: GEIH.

Year	People born in Venezuela (GEIH)	Variation (%)
2005	*36.794*	-
2013	47.693	129.62
2014	80.261	168.28
2015	109.012	135.82
2016	211.071	194.21
2017	413.355	196.18
2018	1’111.225	267.53
2019	1’912.672	172.12

For the year 2005, the value corresponds to the 2005 National Census, and since 2013 this initial value has been projected according to the variations (shifts) coming from the GEIH. In column (3), we observe the variation in the number of Venezuelan migrants residing in Colombia.

in the top 10. Likewise, the cities of the Caribbean Coast (Cartagena, Barranquilla, Soledad, Santa Marta, Maicao, Valledupar, and Sincelejo) also retain a considerable fraction of migrants. While Colombia's capital city and the most populous city (Bogotá) retains the largest share of migrants, Colombia's second and third largest cities (Medellín and Cali) are not in the top 5, as both of them are not close to the Venezuelan border.

On the other hand, the composition of the 2018 Census differs considerably. There, we see that Medellín and Cali are in the top 5 and that Bogotá concentrates a higher proportion of migrants (19.7% compared to 13.5%) and that cities with consolidated migrant networks due to neighborhood ties (such as Cúcuta and Maicao, located on the border) have lower participation. In addition, using 2018, it is observed that foreign people are located mainly in the most populated cities along with their metropolitan areas instead of cities close to the Venezuelan border. In this sense, the proportion of Venezuelan migrants in each municipality must be instrumented to avoid falling into the problem of double causality since the productivity and income of a city will be influenced by migration.

Another source of heterogeneity stemming from migration is its characterization, either by education or income levels. However, although this characterization can be obtained through the GEIH for the study period, it cannot be identified in the 2005 Census, which is why the instrument heterogeneities cannot be constructed. The population without heterogeneous effects will be used instead.

Figure 3: Share of Venezuelan-born in main municipalities: 2005 Census, Migración Colombia

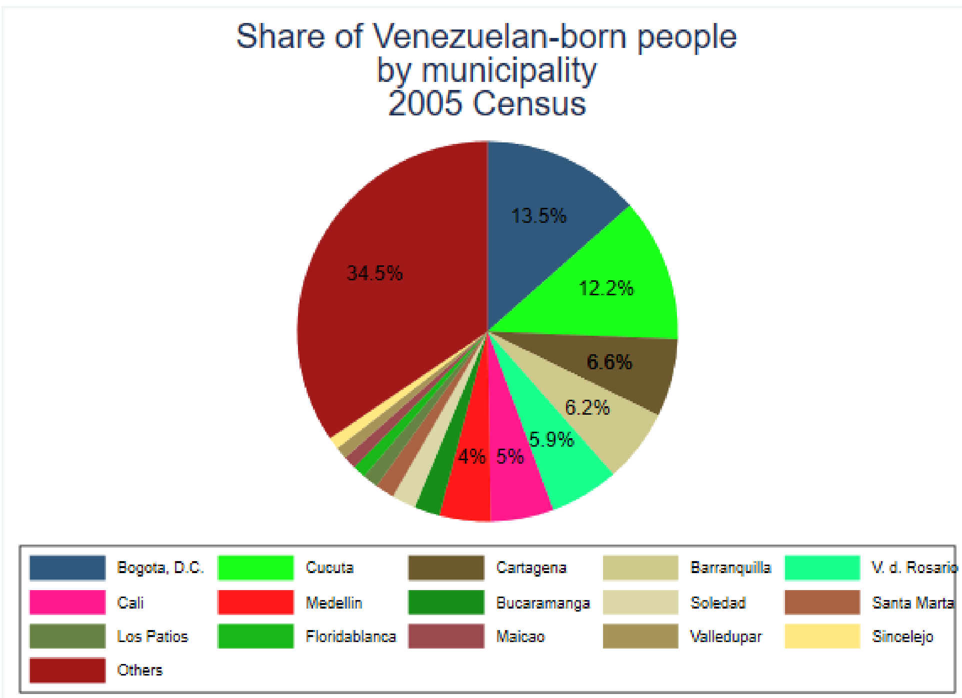
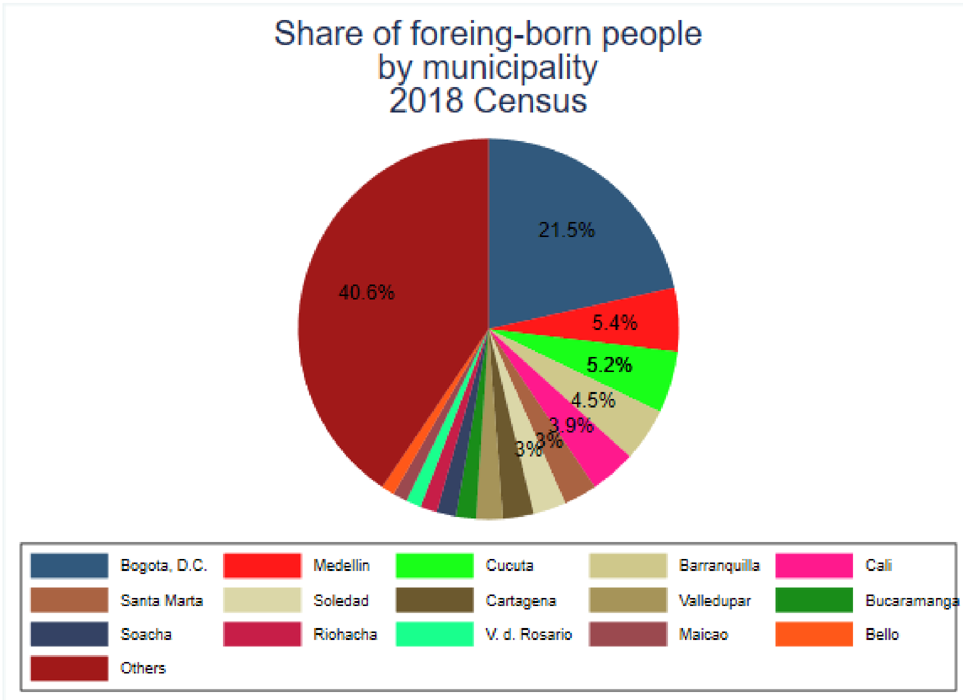


Figure 4: Share of foreign-born in main municipalities: 2018 Census.



4.3.1 Shift-share instrument

Since Venezuelan migrants have not chosen their destinations randomly but rather chose those cities with better economic prospects, an instrumental variables approach is employed to avoid endogeneity, given that the decision of current migrants may be related to current productivity, but not that of past migrants. This methodology has been employed since Altonji and Card (1991). Employing the proportion of Venezuelan inhabitants over the total population in 2005 (the last census before the study period) as a predictor of current migration as shown in Equation (4).

$$\hat{Bartik}_{m,t} = \underbrace{\frac{shareMigracion_{m,2005}}{shareMigracion_{c,2005}}}_{\text{Share}} * \underbrace{\frac{FlujoMigratorio_{ct}}{L_{t-1}}}_{\text{Shift}} \tag{4}$$

This instrument comprises two parts: share (left side of the equation) and shift (right side of the equation). The former is *share* of total Venezuelans residing in municipality j in year t according to the 2005 census, and the latter is the migratory influx received by the whole country according to the GEIH growth rate from Table (2). The left side of Equation (4) corresponds to each city's share for each year, projected to time t using the migratory influx (shift) for the whole country. This component is the variation from one year to another coming from the GEIH. Finally, the total number of residents in Colombia the previous year corresponds to L_{t-1}.

The seminal *enclave* instrumental variable method is prevalent in the migration literature, was first developed by Altonji and Card (1991), and has been employed by several authors such as Del Carpio and M. (2015) and recently, Daly (2021) and Delgado-Prieto (2021). However, Jaeger, Ruist, and Stuhler (2018) argue that this type of past projection in simple enclave instruments is biased in that migration patterns are correlated over time so that the same cities receive migrants. In this regard, Borusyak and Jaravel (2017) discuss Bartik instruments (or Shift-share Instrumental Variables, SSIV) in which identification comes from the quasi-random assignment of shocks. That is weighted averages of the migration, with this weighting being the heterogeneous exposure to the shock. The impact of a set of shocks and units differentially exposed to them will be exploited with exposure measured by weights. This weight can be observed on Equation (4) as the term L_{t-1}.

5 Statistics

Since this investigation uses two different information sources (SIREM and SIIS) to estimate the production function, it is necessary to evaluate the difference between each survey's information. This paper evaluated the following periods: i) SIREM (2013-2015), ii) SIIS (2016-2018), and iii) for the whole period (2013-2018). However, it will consider only the first two options as the data quality differs considerably between them. This exercise will allow us to highlight time heterogeneities as Venezuelan migration in the 2016-2018 period was almost ten-fold compared to 2013-2015.

Table (3) shows the number of firms by municipality. In 2016 there was a decrease in the number in all municipalities, which in turn are the number of firms that must report financial statements to Supersociedades. However, there was an increase in both 2017 and 2018. The municipalities corresponded to the top 34, where more than 50 firms reported their financial statements.

The dependent variable at the city-level stage (aggregated productivity) is the sum of the productivity of all firms located in the city c in a given period t. As it is observed in Table (14). The cities with the highest city-level productivity are located in Andean Region and follow the same pattern similar to the most populated cities in Colombia. Bogotá came in first place with an index of 156,643 in 2013. The following cities are Medellín, Cali, Barranquilla, Bucaramanga and Cartagena. Many other cities are municipalities surrounding the main six cities, for example, Cota (west of Bogotá).

As can be seen, in 2016, all cities suffered a sharp decline in aggregated productivity. However, this is a product of a reduction in data quality in SIIS' financial statements rather than a trajectory in productivity, as observed in Table (15). Bogotá accounts for almost half of total operational revenue and has maintained a similar share after 2016. This pattern is also observed in the rest of the cities. For this reason, the dependent variable will be standardized to render a common magnitude in both periods.

Both in operating income and TFP, we observe in Table (16) similar shares of the top 20 cities with the most firms in the panel. It is observed that Bogota's participation, together with other cities such as Cúcuta, Manizales, and Ibagué has decreased. In contrast, municipalities that are part of the Metropolitan Area have increased, as is the case of Chía, Cota, Funza, La Estrella, and Yumbo.

Table 3: Number of firms by municipality

Number	Municipality	Deparment	2013	2014	2015	2016	2017	2018
1	Armenia	Quindío	94	94	104	90	90	119
2	Barranquilla	Atlántico	1,157	1,134	1,185	962	868	1,148
3	Bogotá D.C.	Bogotá D.C.	14,256	13,513	13,179	10,839	9,157	10,118
4	Bucaramanga	Santander	550	528	555	450	375	558
5	Cali	Valle del Cauca	1,685	1,604	1,614	1,376	1,160	1,524
6	Cartagena de Indias	Bolívar	476	513	530	442	387	449
7	Chía	Cundinamarca	163	173	158	163	140	162
8	Cota	Cundinamarca	315	333	334	325	281	296
9	Cúcuta	Norte de Santander	261	255	272	202	148	187
10	Dosquebradas	Risaralda	69	66	79	72	63	74
11	Envigado	Antioquia	196	189	209	206	175	220
12	Floridablanca	Santander	57	68	79	76	70	100
13	Funza	Cundinamarca	106	126	129	161	165	158
14	Girón	Santander	59	60	56	62	62	72
15	Ibagué	Tolima	195	176	180	142	126	149
16	Itagüí	Antioquia	382	378	372	330	302	342
17	La Estrella	Antioquia	132	141	142	143	125	134
18	Manizales	Caldas	278	257	270	212	196	231
19	Medellín	Antioquia	2,295	2,316	2,470	2,133	1,977	2,429
20	Montería	Córdoba	111	114	124	99	75	138
21	Mosquera	Cundinamarca	102	88	97	109	108	122
22	Neiva	Huila	155	139	143	107	75	106
23	Palmira	Valle del Cauca	142	133	99	95	100	128
24	Pasto	Nariño	76	82	91	79	72	104
25	Pereira	Risaralda	285	276	286	232	232	291
26	Rionegro	Antioquia	88	88	101	99	88	122
27	Sabaneta	Antioquia	188	208	216	170	161	165
28	San Andrés	San Andrés	51	79	150	62	50	56
29	Santa Marta	Magdalena	169	174	184	179	170	229
30	Soacha	Cundinamarca	63	66	58	56	51	55
31	Tocancipá	Cundinamarca	57	64	54	77	77	92
32	Valledupar	Cesar	58	72	75	66	64	87
33	Villavicencio	Meta	174	182	196	174	135	195
34	Yumbo	Valle del Cauca	223	234	202	208	204	236

6 Results

6.1 Production Function (firms)

Variable	SIREM (2013-2015)	SIREM-AT (2013-2015)	SIIS (2016-2018)	SIIS-AT (2016-2018)
Labor	0.212**	0.208*	0.104*	0.158**
	(0.003)	(0.004)	(0.006)	(0.011)
Capital	0.068**	0.072*	0.057*	0.059**
	(0.006)	(0.007)	(0.006)	(0.007)
Materials	0.024**	0.023*	0.055*	0.033*
	(0.003)	(0.004)	(0.015)	(0.016)
Observations	41065	36109	8475	4325
Firms	30357	18709	3461	2287

Table 4: Production Function estimates. AT: not allowing free entry and exit of firms.

The first group of analyses estimated the production function estimates for the three production factors: labor, capital, and material. Table (4) contains the results for each specification. The column *SIREM* indicates that only the Production Function (PF) was estimated only for the survey *SIREM* (i.e., 2013-2015 period), while the column SIIS indicates the complementary period (i.e., 2016-2018 period). The suffix AT indicates an estimate correcting attrition; thus, a correction is made for each survey.

From the results in the Table (4) above, we observe no radical changes in the estimates whether or not free entry and exit of firms is considered. At the same time, Wooldridge (2009) model adjusts the estimates as the sample increases and corrects for entry and exit. In addition, it should be added that firms supervised by *Supersociedades* and do not appear in a single year are not necessarily out of function since financial and non-financial risks are also part of the criteria to be supervised. The results also show that labor is the factor with the highest output elasticity, followed by capital and intermediate materials, whether attrition is or not considered. As results are consistent under all classifications, the estimates will not be corrected for attrition from now on. Aggregated municipality productivity is the sum of PTFs using the parameters found in columns (1) and (3) for each survey and can be found in Table (14) for the main 34 municipalities. In addition, considering the PTF as the portion of the product that the factors can not explain, a considerable residual is observed under the estimation with *Supersociedades*, as previously found on Casas and Balat (2018).

6.2 Relevance of the instrument

One of the two assumptions of instrumental variables is relevance. The Bartik shift-share instrument overcomes the endogeneity problem and correlates significantly with the endogenous variable. In this sense, the table (5) presents the estimates when the dependent variable is the number of foreign-born people in each municipality, and the regressor is the Bartik instrument. As per F-statistic and its correspondent p-value, we observe that an increase of 1% in the instrument leads to 122.8 more foreigners registered in a municipality with a significance level of 99%.

It is essential to consider that the 2018 National Census does not identify the country of origin of foreigners. Nonetheless, the overwhelming majority of foreigners come from Venezuela and, according to Colombian Statistics Department (DANE, in Spanish), account for 3.4% of the Colombian population. Therefore, this result indicates that the Bartik instrument captures not only where foreign migrants choose to locate but also Venezuelan-born people through the previously established networks.

	Number of foreign-born people	
	(1)	(2)
Bartik	623.624***	122.843***
	(42.107)	(13.963)
Controls	No	Yes
N	421	420
F	219	1242
Prob > F	0.000	0.000

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Foreign-born people as the dependent variable using bartik instrument as the regressor.

6.3 City-level estimation

Once the firm-level productivities are added to the aggregated city-level productivity, we will proceed to evaluate the causal effect of Venezuelan migration. As with the production function, we will carry out the estimation for two different periods. The *SIREM* column indicates the years 2013-2015 whereas the *SIREM-SIIS* column represents 2016-2018. In addition, each threshold considers two iterations: i) no control variables and fixed effects by municipality and year, and ii) with control variables and fixed effects by municipality and year.

Since Colombia has 1103 municipalities, and many of them have few firms in both SIREM and SIIS databases, it is necessary to establish a threshold from which the municipality will be included. This exercise considers that the distribution of firms per municipality is not normal. The bulk of economic activity concentrates on the main cities. Therefore, it will consider the size heterogeneity as it changes the threshold in the estimation. The 70th, 75th, 80th, 85th, 90th, and 95th percentiles, including lower, upper and middle limits. For example, the number 70 as the lower limit means that only the municipalities at the right of 70th percentile are included in the estimation. The medium limit refers to not including the distribution extremes and focusing only on the ten lower percentiles surrounding the percentile 70th, for example.

Table (6) presents the estimates for the Bartik instrumental variable for the 70th percentile as the lower limit. The dependent variable was normalized and put in per capita terms (average productivity per firm). So the coefficient value is given in standard deviations. The average value of the dependent variable is close to zero, and its standard deviation to one. As can be seen, there are no positive effects whether or not control variables are included, even though the parameter is positive. The results go in the same line when using the 70th percentile as the upper and medium limits.

Aggregated city productivity					
SIREM (2013-2015)			SIIS (2016-2018)		
Bartik	0.041 (0.817)	0.098 (0.697)	Bartik	0.001 (0.782)	0.000 (0.994)
Fixed effects	Yes	Yes	Fixed effects	Yes	Yes
Controls	No	Yes	Controls	No	Yes
Observations	210	210	Observations	217	217
Municipalities	75	75	Municipalities	76	76
R^2	0.003	0.057	R^2	0.085	0.097
Mean dep. var.	-0.008	-0.008	Mean dep. var.	0.003	0.003
Std. dev. dep. var.	0.986	0.986	Std. dev. dep. var.	0.994	0.994

Table 6: Second stage instrument: Bartik. **Lower limit**.

Aggregated city productivity					
SIREM (2013-2015)			SIIS (2016-2018)		
Bartik	0.261 (0.849)	0.469 (0.736)	Bartik	0.001 (0.761)	0.001 (0.875)
Fixed effects	Yes	Yes	Fixed effects	Yes	Yes
Controls	No	Yes	Controls	No	Yes
Observations	166	166	Observations	129	129
Municipalities	57	57	Municipalities	46	46
R^2	0.037	0.102	R^2	0.044	0.081
Mean dep. var.	-0.015	-0.015	Mean dep. var.	0.007	0.007
Std. dev. dep. var.	0.987	0.987	Std. dev. dep. var.	1.024	1.024

Table 7: Second stage instrument: Bartik. **Upper limit**.

Aggregated city productivity					
SIREM (2013-2015)			SIIS (2016-2018)		
Bartik	0.013 (0.991)	0.283 (0.814)	Bartik	-0.000 (0.952)	0.000 (0.985)
Fixed effects	Yes	Yes	Fixed effects	Yes	Yes
Controls	No	Yes	Controls	No	Yes
Observations	278	278	Observations	223	223
Municipalities	97	97	Municipalities	81	81
R^2	0.010	0.039	R^2	0.007	0.046
Mean dep. var.	-0.051	-0.051	Mean dep. var.	0.004	0.004
Std. dev. dep. var.	0.971	0.971	Std. dev. dep. var.	1.000	1.000

Table 8: Second stage instrument: Bartik. **Medium limit**.

6.4 Robustness

To account for heterogeneities of city size, different thresholds will be considered, using the lower, upper and medium limits described in the previous subsection. To highlight the comparability between each case, parameter values and its standard errors are shown in graphs as follows for the 70th, 75th, 80th, 85th, 90th, and 95th percentiles: i) lower limit (2013-2015) Graph (5), lower limit (2016-2018) Figure (6), upper limit (2013-2015) Figure (7), upper limit (2016-2018) Figure (8), medium limit (2013-2015) Figure (9) and medium limit (2016-2018) Figure (10).

Using different inclusion intervals allows us to evaluate the effect considering heterogeneous and non-linear effects. Although under the lower limit both in the period 2013-2015 and 2016-2018, it is observed that the standard error is reduced as the limit increases, there are no significant effects, neither in the lower limit and medium limit. It is worth considering that in order to preserve a significant number of municipalities in the estimation, only three thresholds were considered.

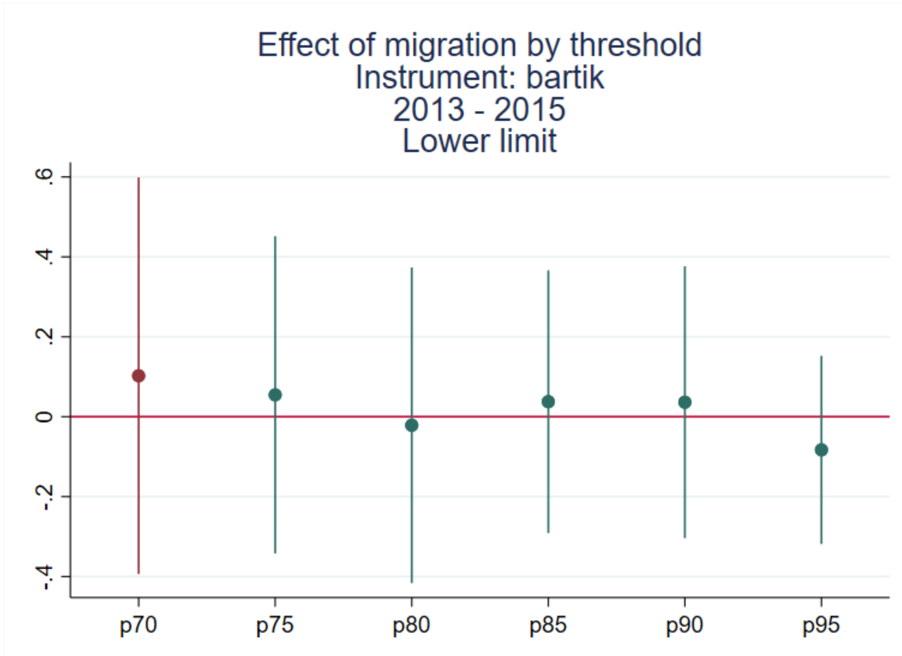


Figure 5: Limit: lower. Period: 2013-2015

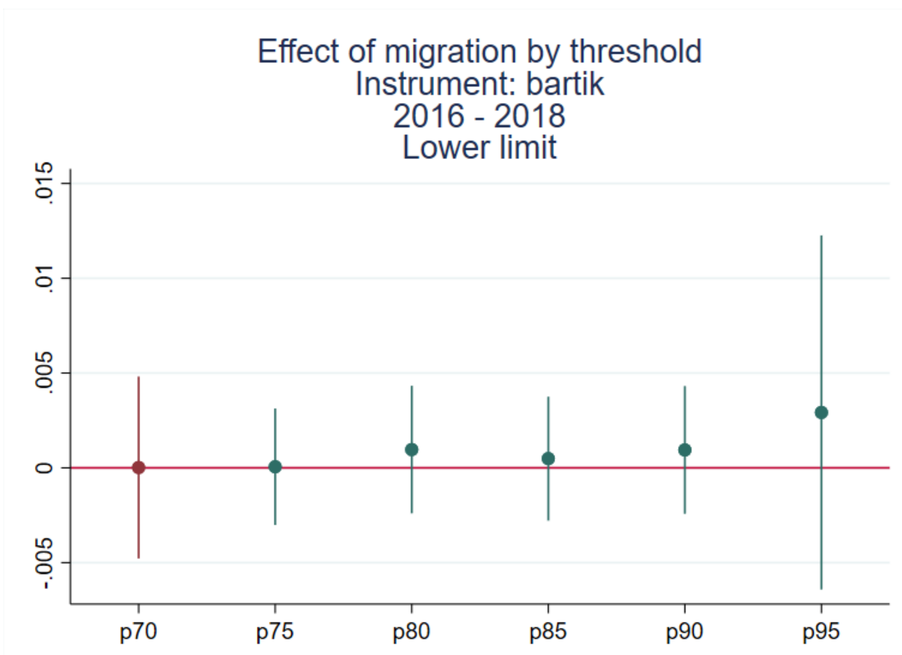


Figure 6: Limit: lower. Period: 2016-2018

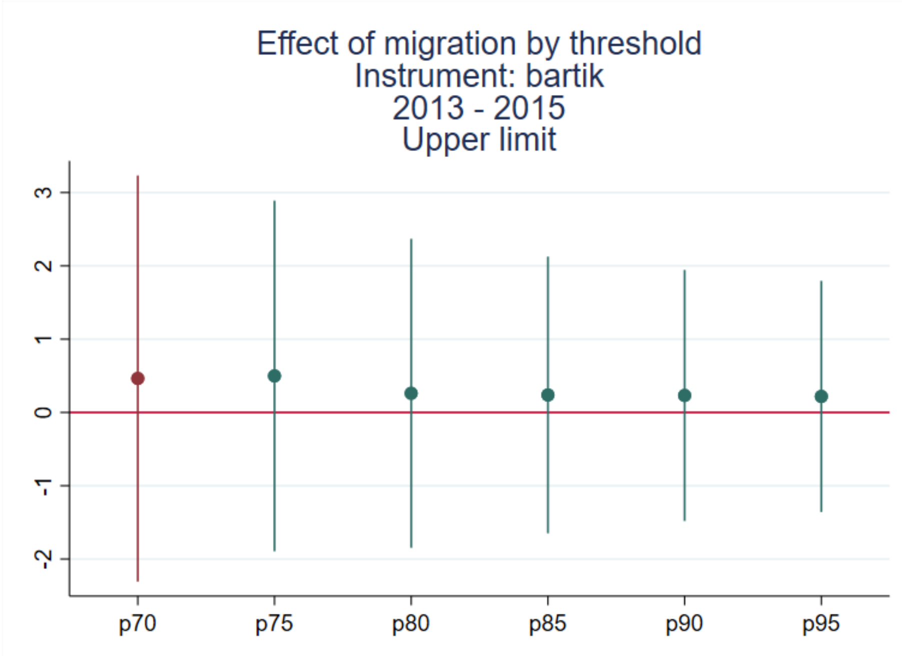


Figure 7: Limit: upper. Period: 2013-2015

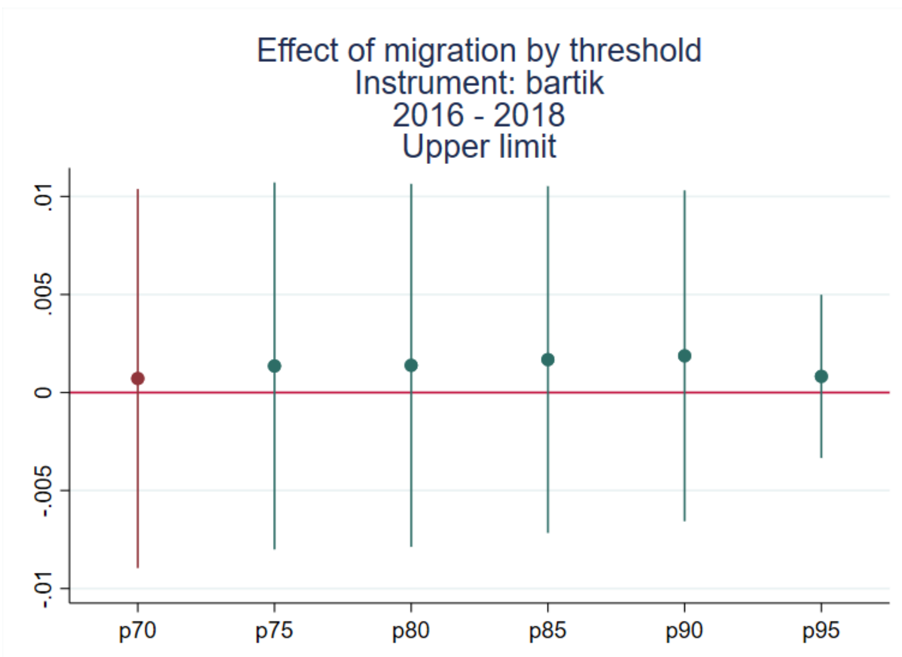


Figure 8: Limit: upper. Period: 2016-2018

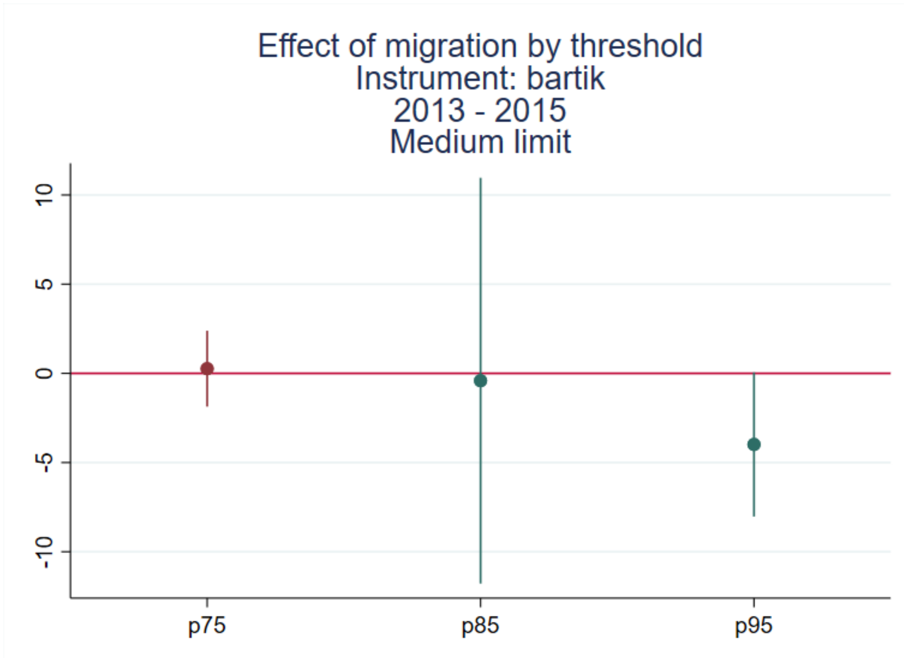


Figure 9: Limit: medium. Period: 2016-2018

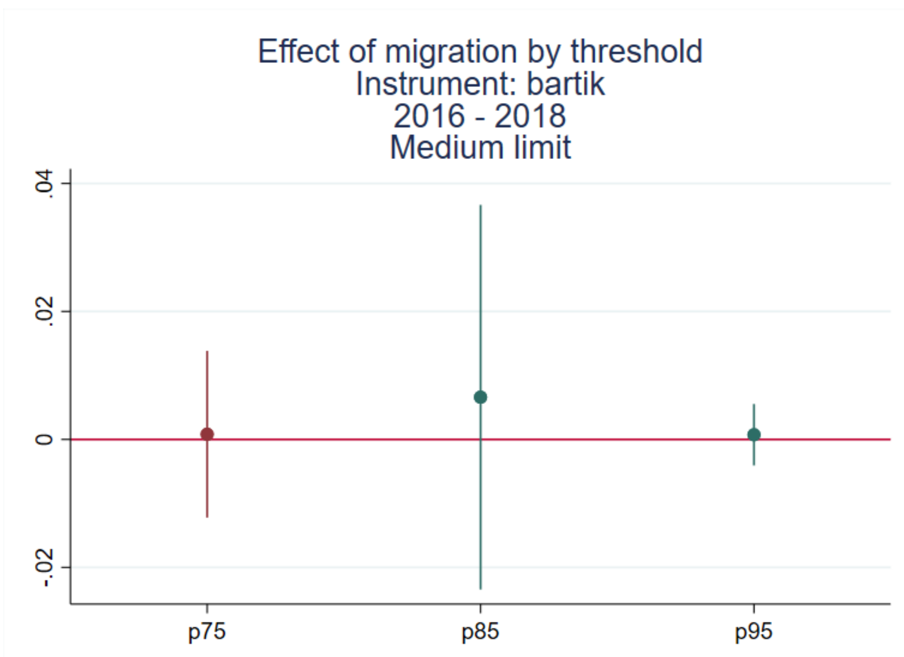


Figure 10: Caption

7 Spillovers

In order to evaluate if Venezuelan migration has affected other significant city outcomes, three different estimations will be carried out considering other dependent variables that go beyond productivity and can also be affected by migration. In the first place, urbanization economies tell us that the size of the city generates economies of scale. According to Sveikauskas, Gowdy, and Funk (1988), the number of city inhabitants is not the correct measure. Instead, he suggests the size of productive activity, which I propose as the size of the operational income of the municipality i in year t. On the other hand, the specialization of the economy in a city also brings economies of scale. For this reason, localization economies are defined as the size of the product of the sectors with high technological requirements in both manufacturing and services, as defined in Table (17) and (18). An additional estimation will be considered when taking the aggregate value for each municipality as the dependent variable. The latter will be referred to as agglomeration effects. Urbanization and localization economies are constructed using SIREM and SIIS, whereas agglomeration economies are obtained from DANE. These variables measure the economies of scale in cities, which are generated by the city's proximity.

These estimations consider the same thresholds (70th, 75th, 80th, 85th, 90th, and 95th percentiles) and limits (lower, upper, medium) as productivity estimations presented in the previous section and are contained in Figure (11) to (28). Regarding the effects of urbanization (size of economic activity), we observe that there are no significant effects at the general level. However, there is an exception for the second period in municipalities between 85th and 95th percentile, where is found a negative effect. However, the pattern holds as no significant effects are found in any other percentile.

Regarding location effects (economic specialization), we observe that migration has a positive effect. In Figure (17), when considering the 70th percentile as the lower bound in the 2013-2015 period, we observe an increase between 0.5 and 1.7 standard variations in aggregate productivity at the municipal level. We see that as the inclusion threshold increases, the mean of the effect is higher until it reaches its maximum level at the 85th percentile. From there, the mean decreases, and considering only the municipalities located to the right of the 95th percentile, we see that the effect ceases to be significant. These results indicate that Venezuelan migration increased specialization in municipalities with fewer firms, while it did not affect the largest cities. This effect is maintained in the 2016-2018 period and is significant for municipalities to the right of the 95th percentile, as shown in Figure 18. Regarding the upper and intermediate limits, no effect is observed, indicating that migration does not lead to increases in the specialization of cities unless it already has a considerable size in the business network.

Finally, for agglomeration effects (value added), we do not find significant effects using the lower or upper limit as inclusion criteria. However, in figures (27) and (27), we observe that when considering the middle boundary, the effect in cities located between the 75th and 85th percentile is positive, indicating that in intermediate cities Venezuelan migration has increased value added through the provision of goods and services.

7.1 Urbanization effects

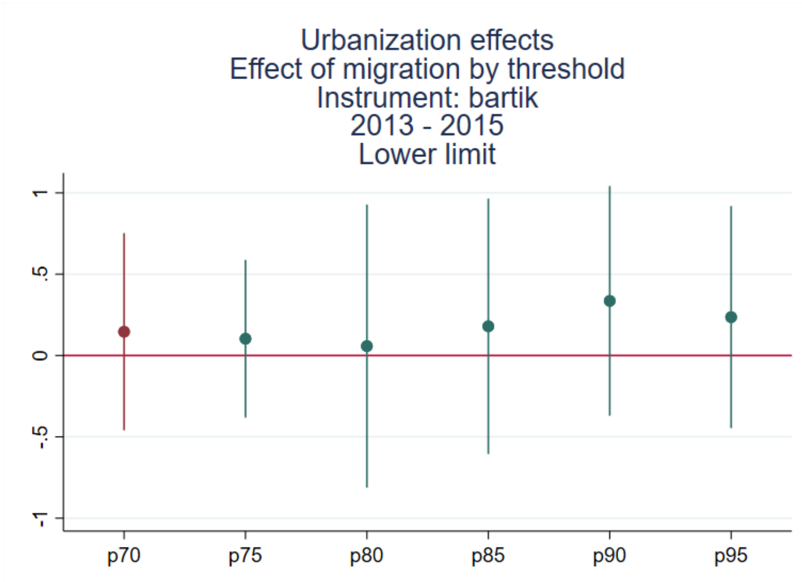


Figure 11: Urbanization effects: Size of the economic activity. Source: SIREM/SIIS. Limit: Lower. Period: 2013-2015.

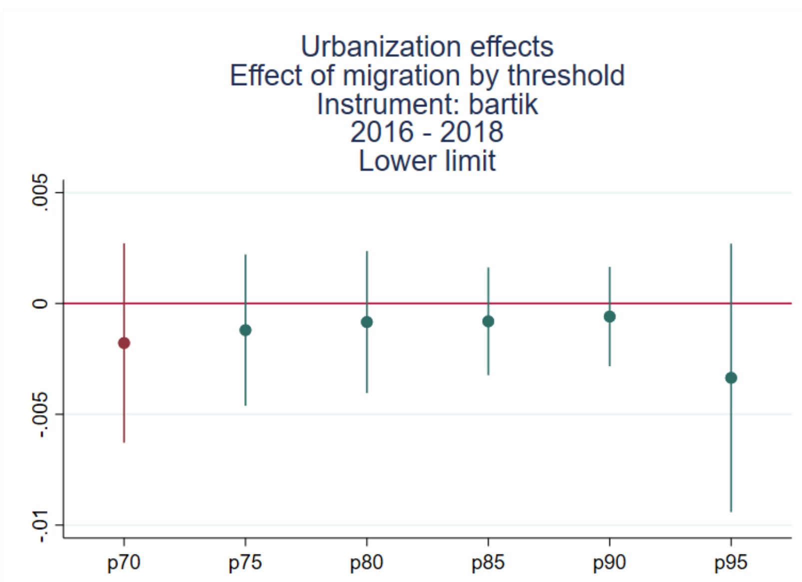


Figure 12: Urbanization effects: Size of the economic activity. Source: SIREM/SIIS. Limit: Lower. Period: 2016-2018.

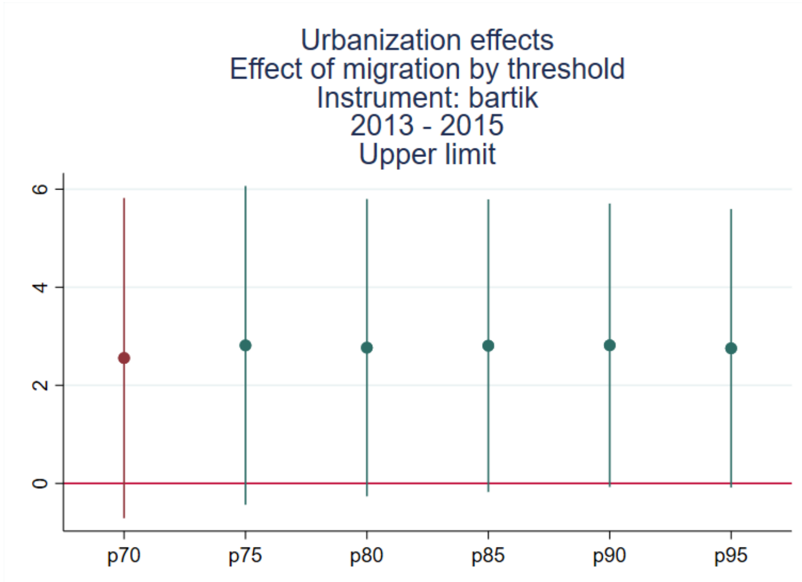


Figure 13: Urbanization effects: Size of the economic activity. Source: SIREM/SIIS. Limit: Upper. Period: 2013-2015.

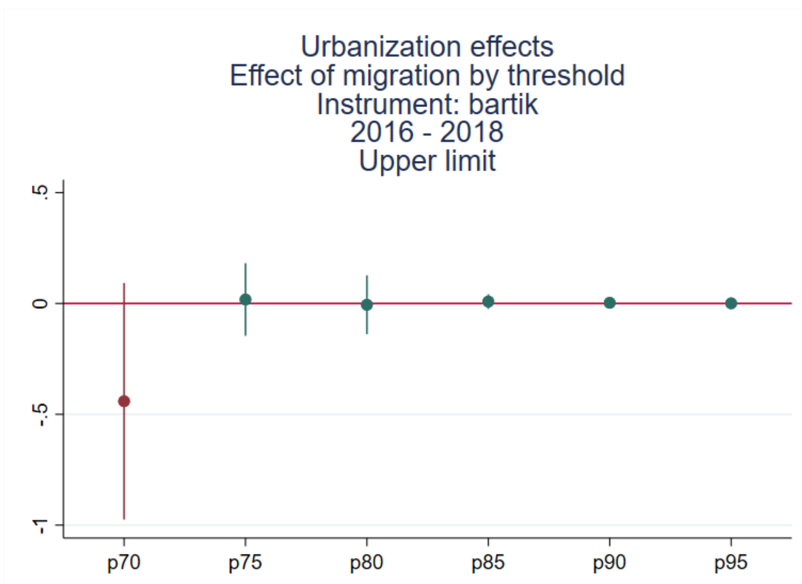


Figure 14: Urbanization effects: Size of the economic activity. Source: SIREM/SIIS. Limit: Upper. Period: 2016-2018.

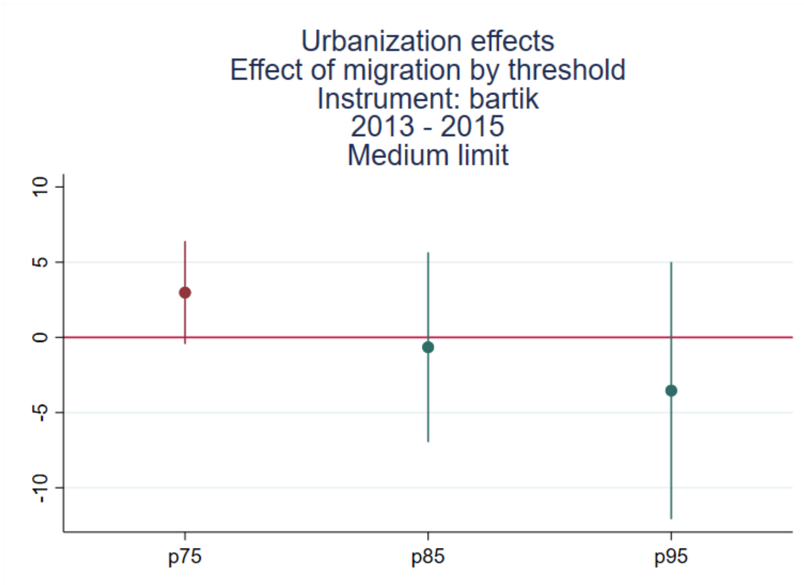


Figure 15: Urbanization effects: Size of the economic activity. Source: SIREM/SIIS. Limit: Medium. Period: 2013-2015.

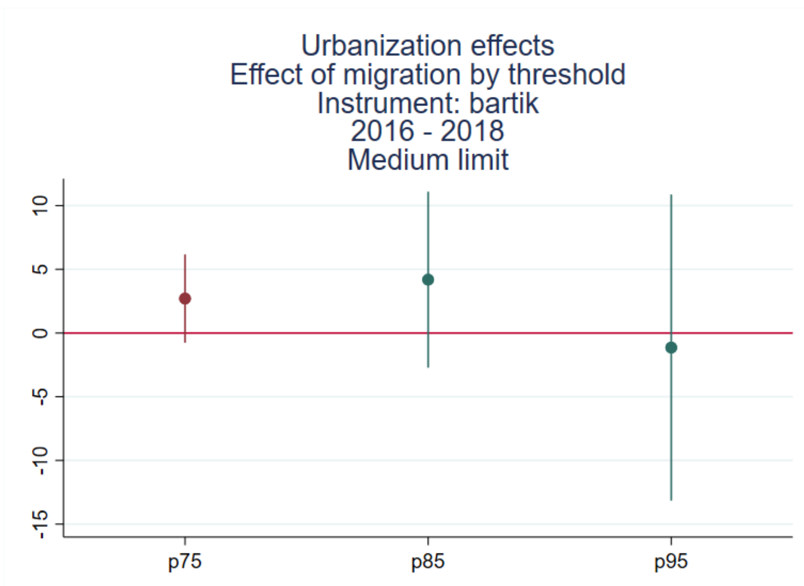


Figure 16: Urbanization effects: Size of the economic activity. Source: SIREM/SIIS. Limit: Medium. Period: 2016-2018.

7.2 Localization effects

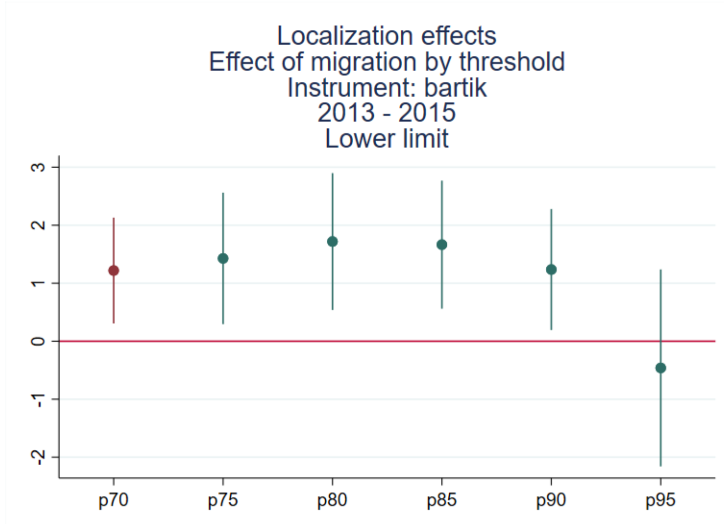


Figure 17: Localization effects: Size of the economic activity in medium-tech and high-tech sectors. Source: SIREM/SIIS. Limit: Lower. Period: 2013-2015.

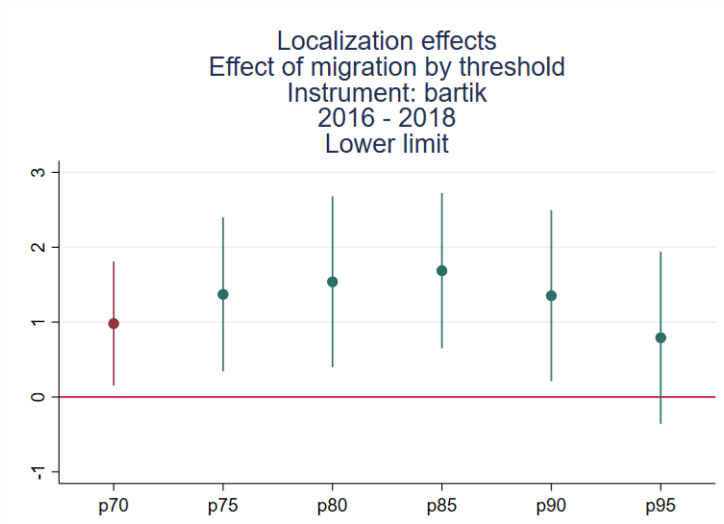


Figure 18: Localization effects: Size of the economic activity in medium-tech and high-tech sectors. Source: SIREM/SIIS. Limit: Lower. Period: 2016-2018.

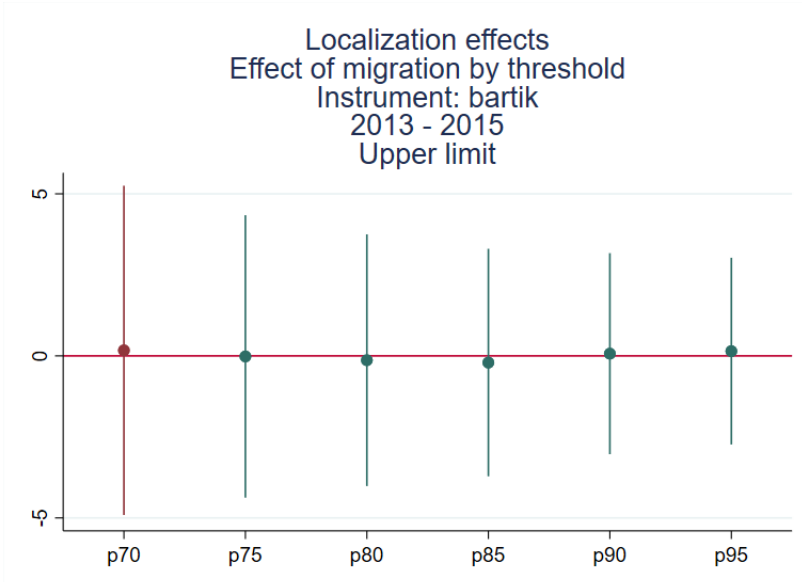


Figure 19: Localization effects: Size of the economic activity in medium-tech and high-tech sectors. Source: SIREM/SIIS. Limit: Upper. Period: 2013-2015.

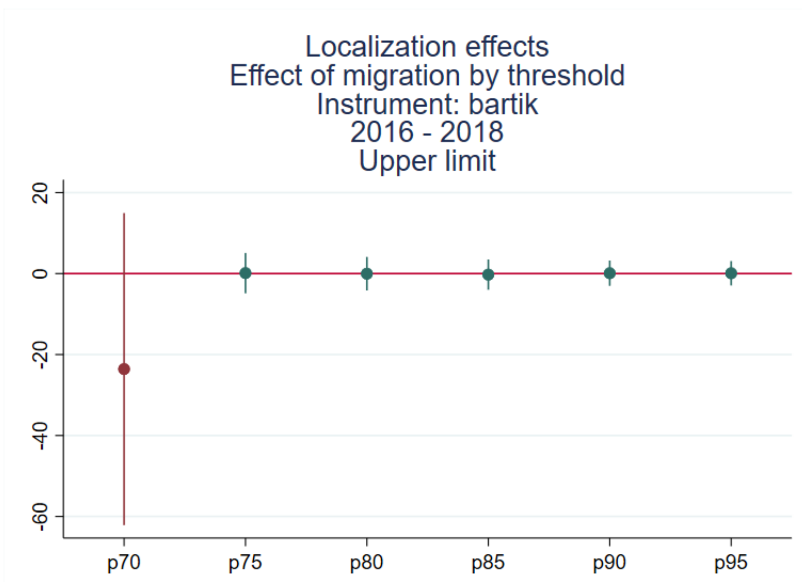


Figure 20: Localization effects: Size of the economic activity in medium-tech and high-tech sectors. Source: SIREM/SIIS. Limit: Upper. Period: 2016-2018.

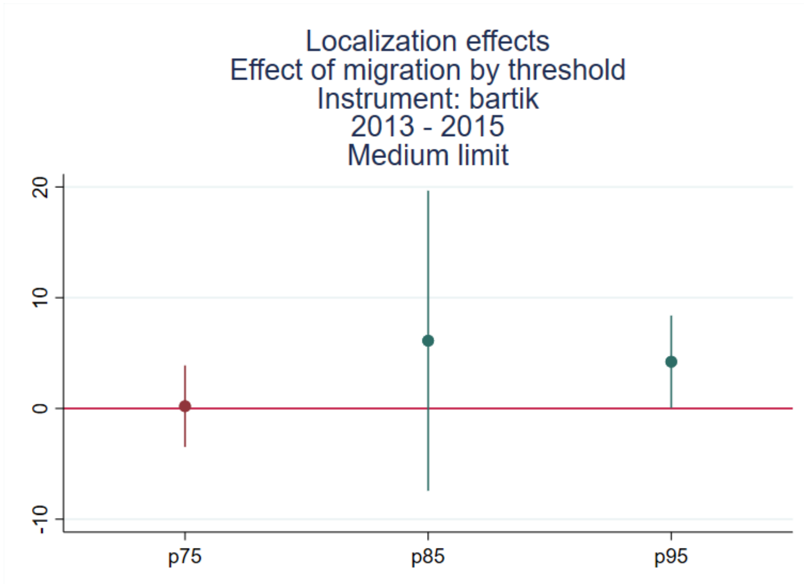


Figure 21: Localization effects: Size of the economic activity in medium-tech and high-tech sectors. Source: SIREM/SIIS. Limit: Medium. Period: 2013-2015.

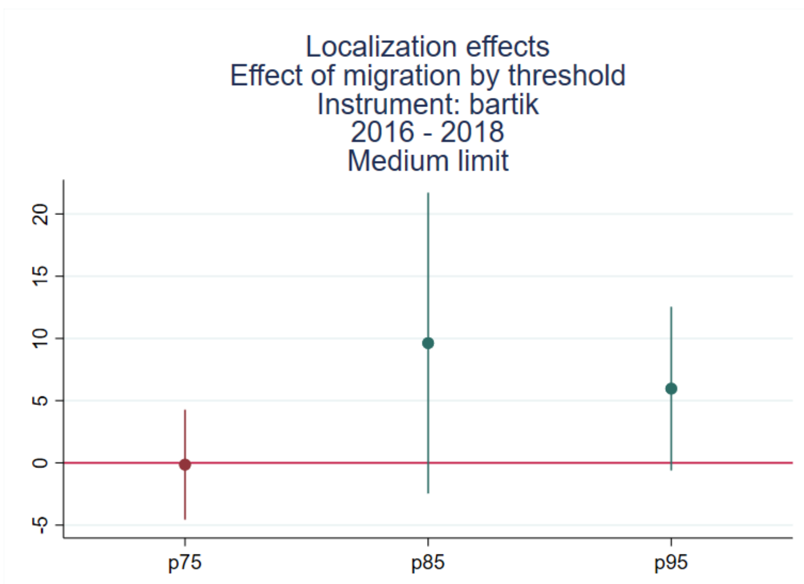


Figure 22: Localization effects: Size of the economic activity in medium-tech and high-tech sectors. Source: SIREM/SIIS. Limit: Medium. Period: 2016-2018.

7.3 Agglomeration effects

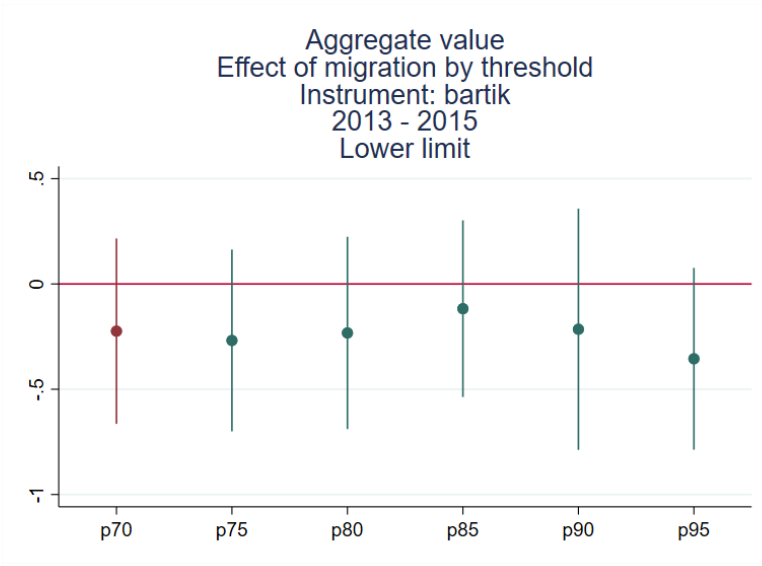


Figure 23: Agglomeration effecs: Value added. Source: DANE. Limit: Lower. Period: 2013-2015.

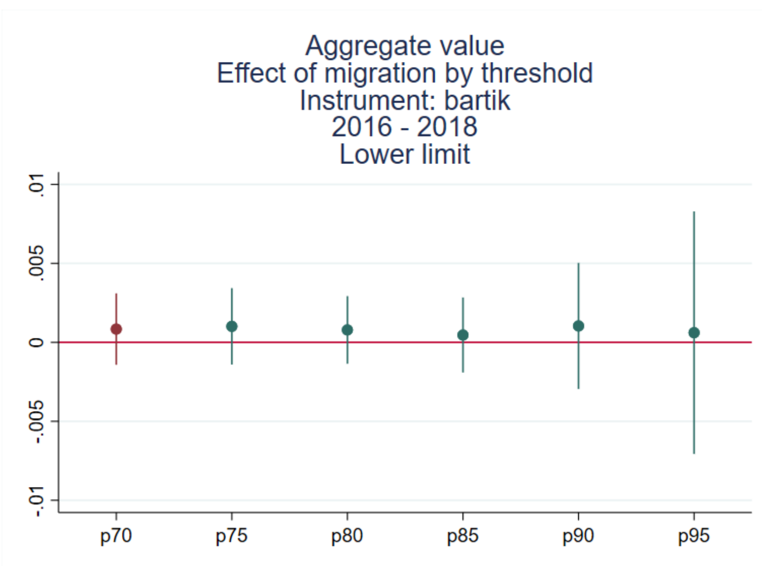


Figure 24: Agglomeration effecs: Value added. Source: DANE. Limit: Lower. Period: 2016-2018.

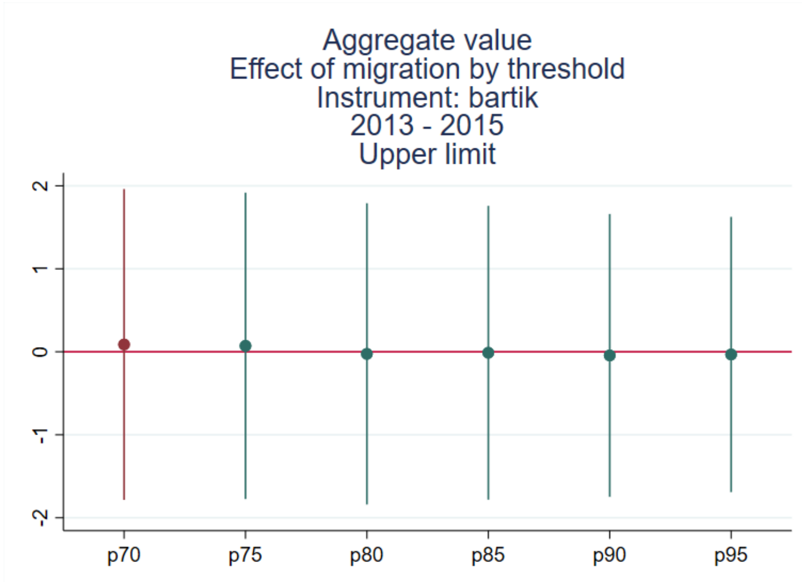


Figure 25: Agglomeration effecs: Value added. Source: DANE. Limit: Upper. Period: 2013-2015.

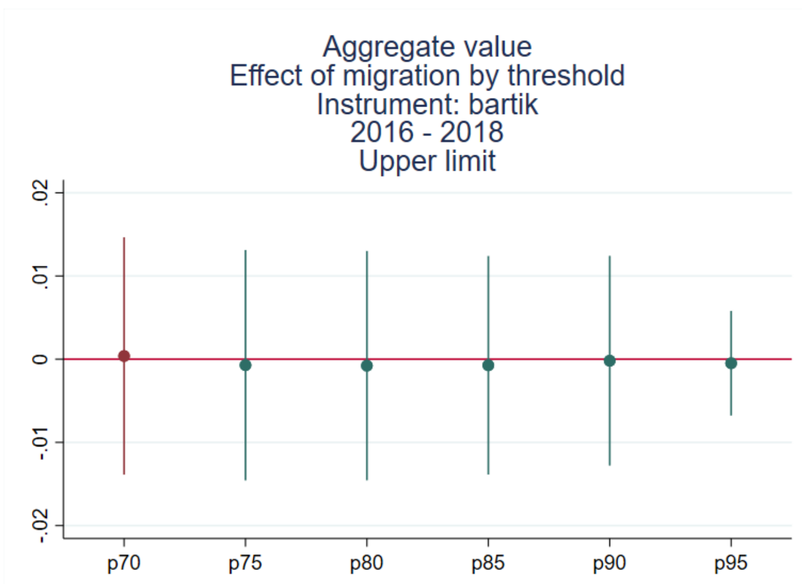


Figure 26: Agglomeration effecs: Value added. Source: DANE. Limit: Upper. Period: 2016-2018.

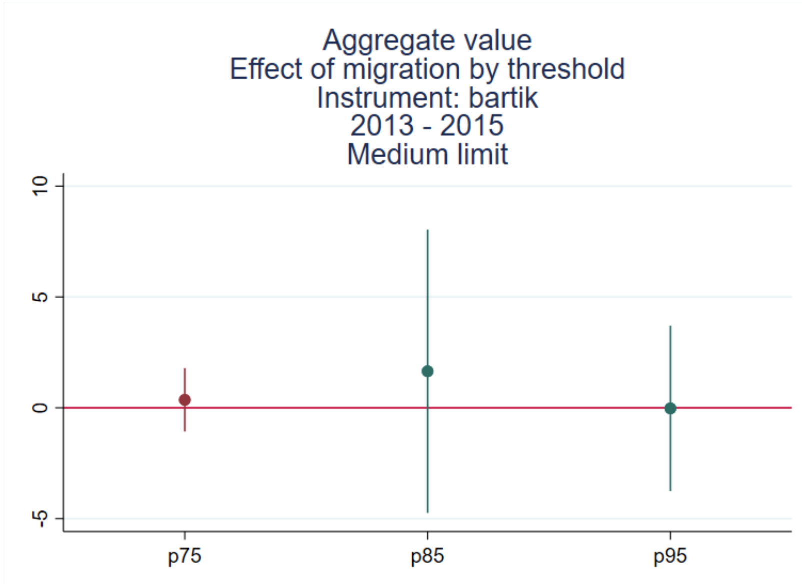


Figure 27: Agglomeration effecs: Value added. Source: DANE. Limit: Medium. Period: 2013-2015.

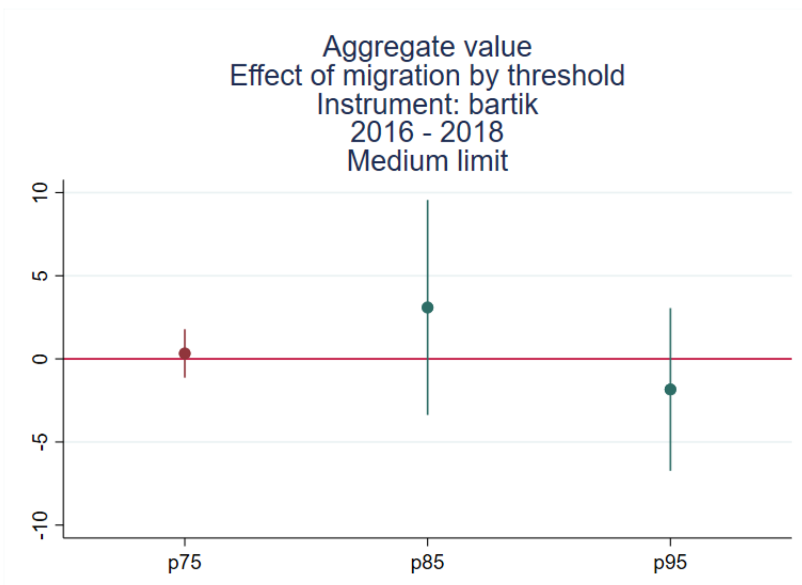


Figure 28: Agglomeration effecs: Value added. Source: DANE. Limit: Medium. Period: 2016-2018.

7.4 Mechanisms

There has been a long discussion concerning migration's absorption by the informal sector (Caruso et al. (2020), Santamaría (2019), Tribin (2020)), which is more than 50% of the labor market in Colombia. Informality has been identified as one of the mechanisms that produce the results presented. Therefore, this paper presents the correlation between migration and productivity. In this sense, it is essential to consider whether migrants locate in cities with greater informality due to the labor market barriers they face or where there is greater formality despite the barriers.

The GEIH contains information on the number of migrants and the proportion of the informal population. In this sense, a panel estimation with fixed effects is performed where the dependent variable is the standardized informality rate, and the regressor variable is the logarithm of the number of Venezuelans. The results in the table 9 indicate significant differences between the location choices of Venezuelan-born migrants from Venezuela and those who were not. The model is robust as per F-statistic. The Venezuelan-born tend to locate where there is less informality, while the non-Venezuelan-born locate regardless of informality levels. This result goes in line with Krugman and Obstfeld (2009) argument. In this sense, informality is not the only source of absorption of migration, so it is paramount to ensure the reduction of barriers in the labor market.

	The migrants were born/lived in Venezuela		
	Born	Five years ago	One year ago
Log. Venezuelan mig.	-0.007*	0.003	-0.002
	(0.086)	(0.543)	(0.612)
Fixed effects	Yes	Yes	Yes
MA* and Departments	36	36	36
R^2	0.510	0.487	0.512
Dep. var. mean	-0.415	-0.405	-0.414
Dep. var. std. deviation	0.217	0.216	0.219
Prob > F	0.000	0.000	0.000

Table 9: Dependent variable: Logarithm of informality in metropolitan areas and departments. (*): Metropolitan Areas.

Discussion

A large number of recent investigations have studied the impact of Venezuelan migration on Colombian outcomes. Most of these investigations focus on the labor market. However, there is no extensive interest in the effect of productivity. This research contributes to migration literature estimating the effect of Venezuelan migration on city-level productivity in Colombia. This investigation is the first work that combines the *SIREM* and *SIIS* databases of *Supersociedades* in the same study.

The estimation presented two endogeneity problems: omitted variable bias for productivity and double causality for the effect of migration on productivity. The strategies to overcome these problems were: i) employing the control function approach to estimate firms' productivity and ii) constructing a shift-share instrument to solve the double causality. The results indicate that city-level productivity at the municipality level has not been positively impacted by Venezuelan migration. Beyond productivity, there is no effect on agglomeration and urbanization economies. However, there is a positive effect in localization economies (specialization) for intermediate and large cities.

These results are consistent with the idea of misallocation of production factors, as pointed out by Martínez and Muñoz (2020). In addition, migrants face hiring barriers, producing a lack of assimilation. Finally, migrants may not adjust to labor demand according to job skills. Since migrants must resort to informal jobs, the misallocation issue is deepened. For these reasons, the south-south migration may not have the same effect reported in the literature, which is mainly based on south-north and north-north migration (Peri (2009), Jahn and Steinhardt (2018)).

However, this study has several limitations. Any official survey provides a characterization of migrant workers. Therefore, considering that Colombia has been receiving international migrants and could be a significant source of work (as the unrestricted mobility of workers was ratified in the 818 Decision from the Comunidad Andina de Naciones - CAN), this could be an incentive to gather this information and improve the possible identification strategies. Secondly, several variables are unavailable at a city level due to the nature of the data, mainly presented at the department and metropolitan area levels. Finally, the composition of migrations given census data can not produce a shift-share instrument. Therefore, other methodologies with worker heterogeneity can be further investigated.

References

Ackerberg, D., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*.

Aguirregabiria, V. (2018). *Empirical industrial organization: Models, methods, and applications.*

Altonji, J. G., & Card, D. (1991, January). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market* (p. 201-234). University of Chicago Press. Retrieved from `http://www.nber.org/chapters/c11773`

Bonilla, L., Morales, L., Hermida, D., & Flórez, L. (2020). El mercado laboral de los inmigrantes y no inmigrantes. evidencia de la crisis venezolana de refugiados. *Borradores de Economía*.

Borjas, G. (1995). The economic benefits from immigration. *The Journal of Economic Perspectives*, *9*(2), 3-22.

Borjas, G., Freeman, R., & Katz, L. (1996). Searching the effect of immigration on the labor market. *Globalization and U.S. Labor Market*.

Borusyak, K., & Jaravel, X. (2017). Revisiting event study designs, with an application to the estimation of the marginal propensity to consume. *Working Paper*.

Caruso, G., Gómez, C., & Mueller, V. (2020). Spillover effects of the venezuelan crisis: migration impacts in colombia. *Oxford Economic Papers*.

Casas, C., & Balat, J. (2018). Firm productivity and cities: The case of colombia. *Borradores de Economía*.

Cárdenas, M., & Mejía, C. (2006). Migraciones internacionales en colombia: ¿qué sabemos? *Fedesarrollo*.

Daly, J. (2021). Effects of venezuelan migration on colombian price level. *Documentos CEDE*.

Del Carpio, X., & M., W. (2015). The impact of syrian refugees on the turkish labor market : The impact of syrians refugees on the turkish labor market. *Policy Research Paper - World Bank*, *http://documents.worldbank.org/curated/en/505471468194980180/The-impact-of-Syrians-refugees-on-the-Turkish-labor-market*.

Delgado-Prieto, L. (2021). Immigration, wages and employment under informal labor markets.

Fabling, R., Maré, D. C., & Stevens, P. (2022). Migration and firm-level productivity. *IZA Institute of Labor Economics*.

Friedberg, R., & Hunt, J. (1995). The impact of immigrants on host country wages, employment and growth. *Journal of Economic Perspectives*.

Glaeser, E. (2011). Cities, productivity, and quality of life. *Science*.

Hanson, G. (2010). International migration and the developing world. *Handbooks in Economics*.

Jaeger, D. A., Ruist, J., & Stuhler, J. (2018). Shift-share instruments and dynamic adjustments: The case of immigration. *Working Paper*.

Jahn, V., & Steinhardt, M. F. (2018). Immigration and new firm formation – evidence from a quasi-experimental setting in germany. *RUHR ECONOMIC PAPERS*.

Jaumotte, F., Koloskova, K., & Saxena, S. C. (2016). Impact of migration on income levels in advanced economies. *International Monetary Fund*.

Krugman, P., & Obstfeld, M. (2009). International economics: Theory and policy. *Boston:Pearson Addison-Wesley*.

Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*.

Martínez, M., & Muñoz, J. (2020). The effect of a labor supply shock on factor's productivity: the case of venezuelan migration in colombia.

Nguyen, T. H. O. (2018). Determinants of firms' total factor productivity inmanufacturing industry in vietnam:an approach of a cross-classified model. *Journal of Asian Business and Economic Studies*.

Olley, G., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *National Bureau of Economic Research*.

Orrenius, P. M., & Zavodny, M. (2012). The economics of u.s. immigration policy. *Journal of Policy Analysis and Management*.

Ottaviano, G. I. P., & Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*.

Peri, G. (2009). The effect of immigration on productivity: evidence from u.s. states. *National Bureau of Economic Research*.

Portes, J. (2019). The economics of migration. *SAGE Journals*.

Pulido, J., & Varon, A. (2020). Misallocation of the immigrant workforce: Aggregate productivity effects for the host country.

Rovigatti, G., & Mollisi, V. (2018). Theory and practice of total-factor productivity estimation: The control function approach using stata. *The Stata Journal*.

Rozo, S., & Vargas, J. (2021). Brothers or invaders? how crisis-driven migrants shape voting behavior. *Vox LACEA - Working Papers*.

Santamaría, J. (2019). Venezuelan exodus: the effect of mass migration on labor market outcomes.

Sveikauskas, L., Gowdy, J., & Funk, M. (1988). Urban productivity: city size or industry size. *Journal of Regional Science*.

Tribin, A. e. a. (2020). Migración desde venezuela en colombia: caracterización del fenómeno y análisis de los efectos macroeconómicos. *Banco de la República*.

Vásquez-DeKartzow, R., Castillo, C., & Durán, L. (2015). Migraciones en países de américa latina. características de la población pediátrica.

Wooldridge, J. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economic Letters*.

Appendix

Control function

Semi-parametric models instrument the endogenous variables to avoid the associated bias. Olley and Pakes (1996) solve a firm's optimization problem and use investment as an instrument of unobserved productivity (ω_{it}). Each year, the firm decides how much to invest and hire in the next period based on its productivity. As per Rovigatti and Mollisi (2018) control function approach is consistent under the following assumptions: i) $i_{it} = f(x_{it}, \omega_{it})$ is the investment policy function, invertible in ω_{it}. Moreover, i_{it} is monotonically increasing in ω_{it}; ii) The state variables (typically capital) evolve according to the investment policy function i_{it}, which is decided at time $t\breve{}1$; and iii) The free variables w_{it} (typically labor inputs and intermediate materials) are non-dynamic, in the sense that their choice at t does not impact future profits, and are chosen at time t after the firm realizes productivity shock. Assuming that unobserved productivity is monotonic, it can be defined as:

$$\omega_{it} = f^{-1}(k_{it}, i_{it})(2)$$

From this transformation, estimation is performed in two steps. First, an approximation of the function from investment and capital using ordinary least squares to estimate the labor parameter l. In the second step, the capital parameter k is estimated assuming that the unobserved productivity depends on the observed productivity in the previous period: And the simultaneity problem is solved because k_{it} was determined by i_{it-1} of the previous period and is therefore exogenous. The above, on the assumption that labor is completely elastic.

In turn, Levinsohn and Petrin (2003) propose to replace the i_{it} investment instrument with intermediate consumption because when the available data presents investment gaps, the assumption that it is monotonic cannot hold. With the new instrument, this assumption continues, and the procedure described by Olley and Pakes (1996) can be continued, while the elastic labor assumption holds.

On the other hand, the Ackerberg, Caves, and Frazer (2015) model excludes the assumption that labor is perfectly adjustable and considers labor to be a function of unobservable productivity it and capital k_{it}, which is determined at time s, being 0<s<1. That is, first, the capital of the next period is selected and then the amount of labor. When we take logarithm from standard Cobb-Douglas functional form and non-observed productivity (ω_{it}) is included in the estimation, as seen in Equation (3), we observe that the β parameter is associated with state variables (capital, w_{it}) while γ is related to the free variables (labor measured by wages or number of workers and materials, x_{it}). However, the term ω_{it}, the productivity not observed by the econometrician, does not have an associated parameter. In addition, Equation (3) assume that productivity shocks affect only free variables (not state variables).

$$y_{it} = \alpha + w_{it}\beta + x_{it}\gamma + \omega_{it} + \varepsilon_{it}(3)$$

Under the seminal approach first developed by Olley and Pakes (1996), productivity evolves as Equation (4), where Ω_{it-1} is the information set at the previous time period and ξ_{it} is the productivity shock, which is uncorrelated with state variables and productivity (Rovigatti & Mollisi, 2018). In Equation (5), Olley and Pakes (1996) find that the optimal solution to a firm's demand of free variables (labor and materials) can be used as a proxy of non-observed productivity. This approach works under the assumption that $f\left(x_{it,}\ \omega_{it}\right)$ is the firm's investment policy function, invertible in ω_{it}, and where capital evolves according to the investment policy function decided at time $t-1$, whereas labor (l_{it}) and materials (m_{it}) are chosen at time t, after the firm realizes its productivity shock. These demands are obtained through the dynamic problem of the firm. Once the proxy is obtained, it is incorporated into the production function.

$$\omega_{it} = E(\omega_{it}|\Omega_{it-1}) + \xi_{it} = E(\omega_{it}|\omega_{it-1}) + \xi_{it} = g(\omega_{it-1}) + \xi_{it}(4)$$

$$\omega_{it} = f^{-1}\left(i_{it,}\ x_{it}\right) = h\left(x_{it,}\ m_{it}\right)(5)$$

Rovigatti and Mollisi (2018) developed a Stata command prodest which estimates production function under different methodologies (e.g., Olley and Pakes and Levinsohn Petrin). Using Wooldridge method, this command performs a consistent estimation within a single step generalized method of moments framework. We prefer Wooldridge (2009) estimation based on many advantages. First of all, its robust standard errors are easily obtained and account for both serial correlation and heteroskedasticity. In addition, this estimation overcomes the potential identification issue highlighted by Ackerberg et al. (2015) in the first stage, such as the correlation between the intermediate input and the error term given the firms' response to technology efficiency shocks.

This method yields econometrically robust standard errors that the instrumental variables may overestimate. This method addresses the Olley and Pakes (1996) and Levinsohn and Petrin (2003) problems by replacing the two-step estimation procedure with a GMM setup. It shows the relevant moment constraints in terms of two equations: both have the same dependent variable (y_{it}) but are characterized by a different set of instruments. There are two equations because the productivity is constructed as follows: i) without imposing any functional form on the control function $\omega_{it} = h(.,.)$; ii) exploiting the markovian nature of productivity (f ...), as observed on equations (6) and (7).The advantages of using this approach are: i) overcoming the potential identification problem highlighted by (**Ackerberg, Caves, Frazer, 2015**) in the first stage and, ii) Robust standard errors are easily obtained, considering both serial correlation and heteroscedasticity.

$$y_{it} = \alpha + w_{it}\beta + x_{it}\gamma + h(x_{it}, m_{it}) + \upsilon_{it} (6)$$

$$y_{it} = \alpha + w_{it}\beta + x_{it}\gamma + f\{h(x_{it-1}, m_{it-1})\} + \eta_{it} (7)$$

The use of simultaneous equations allows us to recover unobserved productivity ($\phi(...)$) in a semi parametric equation, as it is observed in Equation (8), which combines both Equation (6) and Equation (7). On the other hand, y_{it} corresponds to total product, w_{it} contains state variables (capital), x_{it} contains free variables (labor and materials), and both $h(i_{it}, x_{it})$ and $\phi(i_{it}, x_{it})$ indicate the unobserved productivity simultaneously.

$$y_{it} = \alpha + w_{it}\beta + x_{it}\gamma + h\left(i_{it,}\ x_{it}\right) + \varepsilon_{it} = \alpha + w_{it}\beta + x_{it}\gamma + \phi\left(i_{it,}\ x_{it}\right) + \varepsilon_{it}\ (8)$$

Once the productivity proxy is included on the production function, the TFP can be obtained predicting output using each's factor participation and obtaining the difference between observed and estimated output, as Solow residuals are obtained, as is seen in equation (9). Total Factor Productivity is therefore the output that cannot be attributed to the accumulation of capital, labor and materials.

$$\widehat{\omega_{it}} = y_{it} - \widehat{\beta_l} l_{it} - \widehat{\beta_k} k_{it} - \widehat{\beta_m} m_{it} (9)$$

Once the required variables are constructed, we proceed to estimate TFP with the Prodest package. The following options are specified: i) State variable: capital (as it is typical on firm optimization problems, where capital is fixed in the short run but flexible in the long run), ii) Proxy: materials (labor can also be a proxy of productivity, however, we are interested in recover labor participation in the production function) iii) Free: labor, iv) Method: Wooldridge, with GMM and Robinson-Wooldridge variation (using a IV estimation of ACF that employs and IV version of (Robinson, 1988) for estimating output elasticities), v) Polynomial: third degree, and, vi) Repetitions. 50. Per default, the maximum number of iterations to achieve convergence under this command is 10,000. In addition, an attrition correction was considered. However, this estimation do not yield drastically different results.

Production Function

Table 10: Definition and subaccounts included on each production factor.

Item	Definition	Subaccounts
Product	Revenue from ordinary activities before taxes.	Total aggregate.
Capital	Goods of any kind owned by the economic entity, with the intention of using them in the production of goods and services (not intended for sale).	Land, Construction in progress, Machinery and equipment being assembled, Office equipment, Computer and communication equipment, Fleet and transportation equipment, Buildings, Offices and livestock, among others. It also includes the depreciation corresponding to each sub-account.
Labor	Liabilities payable by the economic entity and in favor of employees.	Wages, pensions, severance premiums, benefits and indemnities.
Materials	Includes all those items, materials, supplies, products and renewable and non-renewable resources, to be used in transformation, consumption, rental or sale processes within the ordinary course of business of the economic entity.	Raw materials, products in process, contracts under execution, construction works in progress, among others.

Municipalities

Table 11: Municipalities included on each threshold.

Number	City	Department	50	80	100	150	200
1	Armenia*	Quindío	✓	✓			
2	Barranquilla*	Atlántico	✓	✓	✓	✓	✓
3	Bogotá D.C.*	Bogotá D.C.	✓	✓	✓	✓	✓
4	Bucaramanga*	Santander	✓	✓	✓	✓	✓
5	Cali*	Valle del Cauca	✓	✓	✓	✓	✓
6	Cartagena de Indias*	Bolívar	✓	✓	✓	✓	✓
7	Chía	Cundinamarca	✓	✓	✓		
8	Cota	Cundinamarca	✓	✓	✓	✓	✓
9	Cúcuta*	Norte de Santander	✓	✓	✓		
10	Dosquebradas	Risaralda	✓				
11	Envigado	Antioquia	✓	✓	✓	✓	
12	Floridablanca	Santander	✓				
13	Funza	Cundinamarca	✓	✓	✓		
14	Girón	Santander	✓				
15	Ibagué*	Tolima	✓	✓	✓		
16	Itagüí	Antioquia	✓	✓	✓	✓	✓
17	La Estrella	Antioquia	✓	✓	✓		
18	Manizales*	Caldas	✓	✓	✓	✓	
19	Medellín*	Antioquia	✓	✓	✓	✓	✓
20	Montería*	Córdoba	✓				
21	Mosquera	Cundinamarca	✓	✓			
22	Neiva*	Huila	✓				
23	Palmira	Valle del Cauca	✓	✓			
24	Pasto*	Nariño	✓				
25	Pereira*	Risaralda	✓	✓	✓	✓	✓
26	Rionegro	Antioquia	✓	✓			
27	Sabaneta	Antioquia	✓	✓	✓	✓	
28	San Andrés*	San Andrés	✓				
29	Santa Marta*	Magdalena	✓	✓	✓	✓	
30	Soacha	Cundinamarca	✓				
31	Tocancipá	Cundinamarca	✓				
32	Valledupar*	Cesar	✓				
33	Villavicencio*	Meta	✓	✓	✓		
34	Yumbo	Valle del Cauca	✓	✓	✓	✓	✓

Table 12: Descriptive statistics by municipality. Control variables.

Number	**City / Unit**	**Year**	**Venezuelans** (#)	**Population** (#)	**Distance to capital kms**	**Homicides (per 100inhabitants)**	**Theft (per 100inhabitants)**
1	Armenia	2013	0	4383	21	89	0
2	Barranquilla	2013	2094	1206946	0	27	407
3	Bogotá D.C.	2013	4513	7674366	0	17	359
4	Bucaramanga	2013	729	526827	0	19	575
5	Cali	2013	1790	2319684	0	86	405
6	Cartagena de Indias	2013	2221	978600	0	29	117
7	Chía	2013	105	120719	87	4	31
8	Cota	2013	16	23897	76	9	47
9	Cúcuta	2013	4081	637302	0	42	222
10	Dosquebradas	2013	85	194976	10	31	256
11	Envigado	2013	247	212283	16	14	83
12	Floridablanca	2013	374	263908	9	12	330
13	Funza	2013	19	72566	67	17	63
14	Girón	2013	172	170771	21	15	327
15	Ibagué	2013	141	542876	0	17	392
16	Itagüí	2013	207	261662	12	34	224
17	La Estrella	2013	21	60388	15	67	194
18	Manizales	2013	84	393167	0	29	463
19	Medellín	2013	1456	2417325	0	39	148
20	Montería	2013	174	428579	0	24	298
21	Mosquera	2013	19	78658	60	17	50
22	Neiva	2013	42	337848	0	33	592
23	Palmira	2013	205	300707	39	63	316
24	Pasto	2013	26	428890	0	26	532
25	Pereira	2013	293	464719	0	40	393
26	Rionegro	2013	36	116289	30	33	31
27	Sabaneta	2013	48	50444	16	18	113
28	San Andrés	2013	21	70069	0	27	276
29	Santa Marta	2013	544	469066	0	42	250
30	Soacha	2013	88	488995	51	47	84
31	Tocancipá	2013	1	30326	105	17	47
32	Valledupar	2013	347	433242	0	23	232
33	Villavicencio	2013	154	463121	0	30	461
34	Yumbo	2013	36	111753	25	65	353

Table 13: Descriptive statistics by municipality. Control variables

Number	City / Unit	Year	Tax revenue (per person)	SGP transferences (per person)	Localization revenue (total)	Urbanization revenue (total)	Ratio Localization / Urbanization (%)
1	Armenia	2013	$ 0.28	$ 0.42	536	1301	41.20
2	Barranquilla	2013	$ 0.49	$ 0.43	8482	15479	54.79
3	Bogotá D.C.	2013	$ 0.71	$ 0.28	110056	187895	58.57
4	Bucaramanga	2013	$ 0.45	$ 0.36	3642	7427	49.03
5	Cali	2013	$ 0.34	$ 0.30	11863	22154	53.55
6	Cartagena de Indias	2013	$ 0.47	$ 0.46	3545	6384	55.53
7	Chía	2013	$ 0.51	$ 0.28	1166	2201	52.99
8	Cota	2013	$ 2.69	$ 0.19	3270	4827	67.74
9	Cúcuta	2013	$ 0.19	$ 0.49	1827	3367	54.28
10	Dosquebradas	2013	$ 0.22	$ 0.38	492	1036	47.53
11	Envigado	2013	$ 0.72	$ 0.19	1123	2839	39.57
12	Floridablanca	2013	$ 0.29	$ 0.30	301	759	39.69
13	Funza	2013	$ 0.46	$ 0.10	832	1633	50.94
14	Girón	2013	$ 0.24	$ 0.34	470	899	52.23
15	Ibagué	2013	$ 0.21	$ 0.43	1309	2604	50.27
16	Itagüí	2013	$ 0.65	$ 0.30	2661	5726	46.47
17	La Estrella	2013	$ 0.53	$ 0.10	696	1980	35.14
18	Manizales	2013	$ 0.27	$ 0.38	1904	3686	51.65
19	Medellín	2013	$ 0.51	$ 0.34	16116	30880	52.19
20	Montería	2013	$ 0.17	$ 0.67	736	1530	48.09
21	Mosquera	2013	$ 0.01	$ 0.53	800	1558	51.38
22	Neiva	2013	$ 0.28	$ 0.48	843	2228	37.86
23	Palmira	2013	$ 0.35	$ 0.38	838	1941	43.19
24	Pasto	2013	$ 0.17	$ 0.49	649	1132	57.37
25	Pereira	2013	$ 0.38	$ 0.43	2122	3914	54.23
26	Rionegro	2013	$ 0.11	$ 0.36	512	1245	41.16
27	Sabaneta	2013	$ 1.07	$ 0.31	1039	2699	38.50
28	San Andrés	2013		$ 0.52	193	705	27.42
29	Santa Marta	2013	$ 0.24	$ 0.51	764	2242	34.08
30	Soacha	2013	$ 0.14	$ 0.36	364	979	37.22
31	Tocancipá	2013	$ 1.66	$ 0.15	312	807	38.70
32	Valledupar	2013	$ 0.20	$ 0.53	348	818	42.58
33	Villavicencio	2013	$ 0.26	$ 0.45	1012	2373	42.65
34	Yumbo	2013	$ 1.04	$ 0.15	1468	3290	44.62

Table 14: City-level productivity.

Number	**Municipality**	**Deparment**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
1	Armenia	Quindío	1081.8	1108.7	1237.6	31.7	32.9	32.3
2	Barranquilla	Atlántico	12985.2	12804.3	13148.9	1043.0	1327.7	1212.7
3	Bogotá D.C.	Bogotá D.C.	156643.8	150437.0	142368.7	19110.0	21741.5	21679.2
4	Bucaramanga	Santander	6184.4	6142.3	6325.2	548.7	636.7	707.3
5	Cali	Valle del Cauca	18556.9	17824.7	17427.0	2665.8	2929.3	3032.3
6	Cartagena de Indias	Bolívar	5296.7	5655.7	5598.8	493.5	720.9	582.6
7	Chía	Cundinamarca	1831.6	1918.9	1647.8	420.5	498.7	513.1
8	Cota	Cundinamarca	3927.5	4156.7	4017.1	840.0	933.7	850.9
9	Cúcuta	Norte de Santander	2815.8	2919.6	3022.3	111.0	125.5	127.1
10	Dosquebradas	Risaralda	843.9	809.7	969.9	197.4	184.9	158.1
11	Envigado	Antioquia	2328.8	2226.1	2310.9	487.9	545.3	586.3
12	Floridablanca	Santander	623.6	761.0	852.7	107.0	156.6	133.7
13	Funza	Cundinamarca	1325.5	1564.9	1581.3	357.0	481.7	425.4
14	Girón	Santander	731.0	748.4	659.9	119.1	135.7	145.6
15	Ibagué	Tolima	2150.7	1942.2	2012.3	56.1	99.6	86.9
16	Itagüí	Antioquia	4698.3	4723.6	4615.3	623.6	767.4	719.5
17	La Estrella	Antioquia	1607.3	1690.7	1698.3	310.8	351.3	316.5
18	Manizales	Caldas	3083.3	2934.8	2956.4	218.3	301.3	296.7
19	Medellín	Antioquia	25679.9	25987.2	26959.2	3570.4	4171.8	4402.2
20	Montería	Córdoba	1274.1	1320.4	1419.2	0.0	0.0	13.7
21	Mosquera	Cundinamarca	1274.2	1058.7	1123.7	167.9	282.3	351.5
22	Neiva	Huila	1857.0	1656.8	1655.7	56.5	57.0	55.8
23	Palmira	Valle del Cauca	1605.8	1575.2	1072.5	503.0	688.6	679.0
24	Pasto	Nariño	929.3	1021.5	1102.2	0.0	11.0	0.0
25	Pereira	Risaralda	3253.7	3157.9	3169.2	342.0	429.3	439.5
26	Rionegro	Antioquia	1018.7	1076.3	1190.2	279.7	304.3	300.0
27	Sabaneta	Antioquia	2203.8	2509.0	2558.7	237.4	295.1	260.5
28	San Andrés	San Andrés	586.8	922.9	1689.7	13.8	28.9	15.3
29	Santa Marta	Magdalena	1834.1	1902.4	1968.4	538.0	609.8	586.9
30	Soacha	Cundinamarca	798.8	810.9	717.3	81.4	113.1	118.6
31	Tocancipá	Cundinamarca	656.3	825.8	649.2	303.6	387.6	378.9
32	Valledupar	Cesar	679.0	801.1	894.2	39.4	54.8	56.5
33	Villavicencio	Meta	1969.0	2138.2	2312.0	69.5	68.1	70.2
34	Yumbo	Valle del Cauca	2680.5	2836.1	2374.2	627.6	865.5	848.2

Table 15: Share in national operational revenue for each city.

Number	**Municipality**	**Deparment**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
1	Armenia	Quindío	0.37	0.38	0.44	0.08	0.07	0.07
2	Barranquilla	Atlántico	4.36	4.37	4.58	2.82	3.03	2.80
3	Bogotá D.C.	Bogotá D.C.	52.91	51.66	49.99	51.55	49.69	49.67
4	Bucaramanga	Santander	2.09	2.11	2.24	1.48	1.45	1.61
5	Cali	Valle del Cauca	6.24	6.10	6.11	7.17	6.68	6.90
6	Cartagena de Indias	Bolívar	1.80	1.95	1.97	1.33	1.64	1.34
7	Chía	Cundinamarca	0.62	0.66	0.58	1.14	1.15	1.19
8	Cota	Cundinamarca	1.36	1.47	1.44	2.29	2.18	1.98
9	Cúcuta	Norte de Santander	0.95	0.99	1.06	0.30	0.29	0.29
10	Dosquebradas	Risaralda	0.29	0.29	0.35	0.53	0.42	0.37
11	Envigado	Antioquia	0.80	0.78	0.82	1.32	1.27	1.34
12	Floridablanca	Santander	0.21	0.27	0.31	0.28	0.36	0.31
13	Funza	Cundinamarca	0.46	0.55	0.57	0.99	1.10	0.99
14	Girón	Santander	0.25	0.26	0.24	0.33	0.32	0.35
15	Ibagué	Tolima	0.73	0.67	0.72	0.15	0.24	0.20
16	Itagüí	Antioquia	1.61	1.64	1.64	1.70	1.78	1.67
17	La Estrella	Antioquia	0.56	0.59	0.61	0.85	0.81	0.74
18	Manizales	Caldas	1.04	1.00	1.04	0.60	0.70	0.67
19	Medellín	Antioquia	8.70	8.91	9.42	9.60	9.56	10.04
20	Montería	Córdoba	0.43	0.45	0.50			0.03
21	Mosquera	Cundinamarca	0.44	0.37	0.40	0.45	0.65	0.81
22	Neiva	Huila	0.63	0.57	0.59	0.16	0.13	0.13
23	Palmira	Valle del Cauca	0.55	0.55	0.39	1.35	1.58	1.56
24	Pasto	Nariño	0.32	0.36	0.40		0.02	
25	Pereira	Risaralda	1.10	1.09	1.12	0.94	1.02	1.04
26	Rionegro	Antioquia	0.35	0.38	0.42	0.74	0.69	0.68
27	Sabaneta	Antioquia	0.76	0.88	0.91	0.65	0.67	0.57
28	San Andrés	San Andrés	0.20	0.31	0.58	0.03	0.07	0.04
29	Santa Marta	Magdalena	0.63	0.67	0.71	1.44	1.39	1.36
30	Soacha	Cundinamarca	0.28	0.29	0.26	0.23	0.27	0.27
31	Tocancipá	Cundinamarca	0.23	0.29	0.23	0.83	0.89	0.89
32	Valledupar	Cesar	0.23	0.28	0.31	0.11	0.13	0.13
33	Villavicencio	Meta	0.67	0.73	0.81	0.18	0.16	0.16
34	Yumbo	Valle del Cauca	0.93	1.01	0.86	1.73	2.05	1.98

Table 16: Share in national PTF for each city.

Number	Municipality	Deparment	2013	2014	2015	2016	2017	2018
1	Armenia	Quindío	0.37	0.38	0.44	0.09	0.08	0.07
2	Barranquilla	Atlántico	4.40	4.41	4.63	2.82	3.04	2.78
3	Bogotá D.C.	Bogotá D.C.	53.02	51.78	50.09	51.62	49.84	49.76
4	Bucaramanga	Santander	2.09	2.11	2.23	1.48	1.46	1.62
5	Cali	Valle del Cauca	6.28	6.13	6.13	7.20	6.72	6.96
6	Cartagena de Indias	Bolívar	1.79	1.95	1.97	1.33	1.65	1.34
7	Chía	Cundinamarca	0.62	0.66	0.58	1.14	1.14	1.18
8	Cota	Cundinamarca	1.33	1.43	1.41	2.27	2.14	1.95
9	Cúcuta	Norte de Santander	0.95	1.00	1.06	0.30	0.29	0.29
10	Dosquebradas	Risaralda	0.29	0.28	0.34	0.53	0.42	0.36
11	Envigado	Antioquia	0.79	0.77	0.81	1.32	1.25	1.35
12	Floridablanca	Santander	0.21	0.26	0.30	0.29	0.36	0.31
13	Funza	Cundinamarca	0.45	0.54	0.56	0.96	1.10	0.98
14	Girón	Santander	0.25	0.26	0.23	0.32	0.31	0.33
15	Ibagué	Tolima	0.73	0.67	0.71	0.15	0.23	0.20
16	Itagüí	Antioquia	1.59	1.63	1.62	1.68	1.76	1.65
17	La Estrella	Antioquia	0.54	0.58	0.60	0.84	0.81	0.73
18	Manizales	Caldas	1.04	1.01	1.04	0.59	0.69	0.68
19	Medellín	Antioquia	8.69	8.94	9.48	9.64	9.56	10.10
20	Montería	Córdoba	0.43	0.45	0.50	0.00	0.00	0.03
21	Mosquera	Cundinamarca	0.43	0.36	0.40	0.45	0.65	0.81
22	Neiva	Huila	0.63	0.57	0.58	0.15	0.13	0.13
23	Palmira	Valle del Cauca	0.54	0.54	0.38	1.36	1.58	1.56
24	Pasto	Nariño	0.31	0.35	0.39	0.00	0.03	0.00
25	Pereira	Risaralda	1.10	1.09	1.11	0.92	0.98	1.01
26	Rionegro	Antioquia	0.34	0.37	0.42	0.76	0.70	0.69
27	Sabaneta	Antioquia	0.75	0.86	0.90	0.64	0.68	0.60
28	San Andrés	San Andrés	0.20	0.32	0.59	0.04	0.07	0.04
29	Santa Marta	Magdalena	0.62	0.65	0.69	1.45	1.40	1.35
30	Soacha	Cundinamarca	0.27	0.28	0.25	0.22	0.26	0.27
31	Tocancipá	Cundinamarca	0.22	0.28	0.23	0.82	0.89	0.87
32	Valledupar	Cesar	0.23	0.28	0.31	0.11	0.13	0.13
33	Villavicencio	Meta	0.67	0.74	0.81	0.19	0.16	0.16
34	Yumbo	Valle del Cauca	0.91	0.98	0.84	1.70	1.98	1.95

High-tech classification

Table 17: Manufacturing subsectors classified as high-tech. Source: Eurostat, the statistical office of the European Union.

Manufacturing industries	NACE Rev. 2 codes	2-digit level
High-technology	21	Manufacture of basic pharmaceutical products and pharmaceutical preparations;
	26	Manufacture of computer, electronic and optical products,
Medium-high-technology	20	Manufacture of chemicals and chemical products; Manufacture of electrical equipment; Manufacture of machinery and equipment n.e.c.
	27 to 30	Manufacture of motor vehicles, trailers and semi-trailers; Manufacture of other transport equipment
Medium-low-technology	19	Manufacture of coke and refined petroleum products;
	22 to 25	Manufacture of rubber and plastic products; Manufacture of other non-metallic mineral products; Manufacture of basic metals; Manufacture of fabricated metals products, excepts machinery and equipment;
	33	Repair and installation of machinery and equipment
Low technology	10 to 18	Manufacture of food products, beverages, tobacco products, textile, wearing apparel, leather and related products, wood and of products of wood, paper and paper products, printing and reproduction of recorded media;
	31 to 32	Manufacture of furniture; Other manufacturing

Table 18: Services subsectors classified as high-tech. Source: Eurostat, the statistical office of the European Union.

Services firms	NACE Rev. 2 codes	2-digit level Knowledge-intensive services (KIS)
Knowledge-Intensive Firms (KIS), High-tech	50 to 51	Water transport; Air transport;
–	58 to 63	Publishing activities; Motion picture, video and television programme production, sound recording and music publish activities; Programming and broadcasting activities; Telecommunications; computer programming, consultancy and related activities; Information service activities (section J);
	64 to 66	Financial and insurance activities (section K);
	69 to 75	Legal and accounting activities; Activities of head offices, management consultancy activities; Architectural and engineering activities, technical testing and analysis; Scientific research and development; Advertising and market research; Other professional, scientific and technical activities; Veterinary activities (section M);
	78	Employment activities;
	80	Security and investigation activities;
	84 to 93	Public administration and defence, compulsory social security (section O); Education (section P), Human health and social work activities (section Q); Arts, entertainment and recreation
Less knowledge-intensive services (LKIS)	45 to 47	Wholesale and retail trade; Repair of motor vehicles and motorcycles (section G);
	49	Land transport and transport via pipelines;
	52 to 53	Warehousing and support activities for transportation; Postal and courier activities;
	55 to 56	Accommodation and food service activities (section I);
	68	Real estate activities (section L);
	77	Rental and leasing activities;
	79	Travel agency, tour operator reservation service and related activities;
	81	Services to buildings and landscape activities;
	82	Office administrative, office support and other business support activities;
	94 to 96	Activities of membership organisation; Repair of computers and personal and household goods; Other personal service activities (section S);
	97 to 99	Activities of households as employers of domestic personnel; Undifferentiated goods- and services-producing activities of private households for own use (section T); Activities of extraterritorial organisations and bodies (section U)