VENEZUELAN MIGRATION AND INFORMAL FIRMS EXPANSION

Universitat Autònoma de Barcelona
Diana Gutiérrez

June 30, 2023

Abstract

Migration literature has focused on how native labor outcomes are affected by migration, concluding null effects. Some offsetting mechanisms must operate to explain those results. Traditional results rely on the labor supply side but less is said from the demand side regarding how firms react when facing a migratory shock. Let along when the informal sector is large as occurs in developing countries. We show evidence for one mechanism in developing cities that is firms react by expanding the number of informal firms *(Extensive Margin)*. By implementing a fixed effects Durbin Spatial Model, we test whether the extensive margin, mostly in the informal sector, could explain the immigrants absorption from Venezuelan across 1.171 neighborhoods in Bogotá-Colombia. Also, existance of urban spatial spillover effects helps intensify the extensive margin. Results confirmed that, on average, a 1% rise in the immigrants density is associated with a rise in the number of formal and informal establishments by km2, close to 1%. Spillover Effects reinforce the extensive margin of informal firms five times more than in the case of the formal units. Spatial heterogeneity analysis is also present as Venezuelan migration is related to promote expansion in the formal establishments location at the central CBD before 2018, afterwards, extensive margin was greater for the infomal ones at the south west periphery of the city.

Keywords: migration, extensive margin, spillover effects, informality.

JEL codes: R23, L11, C21, J46, R3.

1 INTRODUCTION

Migration effects are more noticeable at the urban level than across the entire national economy (Saiz, 2007). Immigrants can shape cities as they cluster in particular areas. Some of the facets that migration influences at the urban level are the labor markets and firms agglomerations (Cuong et al., 2020). The first reference point of inmigrants for jobsearch and entrepeneurship are those centers with a high concentration of employment and population (Su et al., 2018).

Though, cities that migrants face are completelly different between developed and developing cities as the latter implies urban characteristics that are not usually found among developed nations. Informality of the economic tissue is one of the differential predominant characteristic. The CBDs are a mix of formal and informal firms and as the distance to CBD increases the number of informal firms becomes greater. ¿How can migrants influence the shape of firms agglomerations pattern in developing cities?.There are some potential mechanisms; if the cheaper labor from inmigrants is available for native entrepreneurs, they can be motivated to expand their businesses with new establishments; or the migrants themselves can decide to undertake new bussinesses. In either of the two circumstances, the firms agglomeratios expand driven by migration.

The above is named by Olney (2013) the *Extensive Margin*. That is, the excess labor supply induced by immigrants could potentially be absorbed by firms expanding the number of establishments. Rather than the *Intensive Margin*, that is, the expansion of the production at the individual level of the firms. The latter is much more the mechanism of abpsortion in advanced economies.

In the literature most of the works find not significant effects of migration on employment and wages. Nor for the developed economies (Altonji and Card, (1989; 1991); Schoeni, 1997; Card, (2001, 2007); and Card and Lewis, 2007). Neither for the formal and the informal sector in developing economies analysis (Santamaría, 2020; Bahar et al, 2020; Tribín and Uribe et al. 2020; Morales-Zurita et al., 2020; Tribet al, 2020; Howard, 2019). If we take these findings as given null effects could occur because labor demand adjusts to the labor supply shock leaving wages and unemployment unaffected. Then the labor demand also responds to immigration. We follow Olney in what labor demand reacts throught the extensive margin. Hence, it is worth it approaching it to see whether part of the adjustment mechanism in developing countries is also due to the extensive margin.

For the above, the aim of this study is to examine the extensive margin in developing cities, whether migration can influence the agglomeration pattern by firms expanding the number of establishments and whether this expansion occurs in the formal or informal market. We focus on the case of Venezuelan migration because it is the most massive and recent type of migration from and to a developing country. Also, Colombia is the country that most receive Venezuelans in Latin America (32%), followed by Peru (17.39%) and Chile, (9.67%). We evaluate the relationship across 1.171 neighbourhoods in Bogota for the period 2013-2022.

Our central hypothesis its that the extensive margin view is more usefull for explaining immigration efects in developing citites because they lack of mitigation capacity. In developing countries the informal sector is large [1], generally composed by small units with low production capacity (Nicolini and Roig, 2019) that cannot adopt new technologies to raise productivity and accommodate in the long term a migratory shock at fair wages.

Firms individually cannot respond to absorb the labour excess from immigration. The quality of employment is also affected, because informal workers are not covered by social security, either contracts with legal protection without job stability or permanence (Combarnous et al, 2019). Then, the ability of the informal natives to upgrade skills in the face of migratory shocks is very

[1] Around 80% of businesses and labour in developed economies are formal, while in developing countries more than 70% of the employment is based on the informal sector (World Bank,2020)

weak.

All these shortcomings are not observed in developed countries where individual firms adjust to changes in the skill composition of production factors, in such a way, to absorb the oversupply of labor. Especially, those industries unskilled labour intensive (Leamer and Levinsohn, 1995; Bhagwati and Srinivasan, 1983). Firms make technology adjustment such as equipment and technology improvements, hence, the relatively abundant type of labour is more used intensively (Lewis, 2011, 2013; Peri and Sparber, 2009). In Europe and the United States, immigration has a strong positive association with productivity growth, because immigrants specialize in manual-intensive tasks and native workers in communication-intensive tasks (Constant, 2014). This triggers a relative shortage of skilled work that raises their wages, motivating upgrading skills. In the long-term, labour markets can accommodate foreign workers without harming natives (Constant, 2014).

If intensive margin fails in developing economies, ¿what another absorption mechanism can be found for the informal developing economies?. We believe, it is highly likely that developing cities respond to the oversupply of labor by expanding informal firms massively rather than intensively increasing production.

In addition, two elements should be considered along with the central hypothesis. First, the informal sector provides inputs to the formal and viceversa (Ulyssea (2018, 2020) and Hernandez et al, 2011). Second, that interrelationship is transferred into urban space as the CBD is mixed and with distance increases the rate of informal businesses (Hernandez et al, 2011; Posada et al,2019).

Under the previous description, this document incorporates the spatial dimension of the extensive margin (spatial interdependence). Considering their spillover effects across the space and also the spatial structure of the city. These features are integrated into a hybrid model following García-López et al (2020), whose baseline equation is a Spatial Durbin-SDM model type.

Regarding immigrants data used, several sources of information will be combined in a complementary way; the 2018 Population and Housing Census allows to track Venezuelan immigrants to be counted at the block level from 2012-2018. The administrative records of the Register of Venezuelan Immigrants (RUVM in Spanish) with records of population who required social services (SISBEN database), allow to track Venezuelan immigrants from 2019 to 2022. Here, we defined immigrants as those who were born in Venezuela and arrived yearly for first time into the country. All individual information is aggregated at 1.171 neighborhood level.

In the case of firms, it will be applied the labor-approach definition of International Labor Organization for informality, which defines informal firms as those who do not pay to their employees, pensions and health as part of the salary. Data for individual formal firms is obtained from administrative records of the Social Security Payment Contributions Registry 2018-2022 (named PILA in Spanish) and Chamber of Commerce for the period 2013-2022. Also, it is used two Censuses of Economic Units (2017;202) to track the total number of existing firms in the city from the question “what year started operations”. Regarding informal firms, a measure of “potential informality” is built as the subtraction between the total number of units found in Censuses and the formal units registered in the administrative records. All individual firms are aggregated at 1.171 neighborhood level.

Results suggest that the extensive margin operates for informal and formal fims in Bogotá. That is, immigration is associated with an increase in the number of establishments within a city. A 10 percent increase in the density of immigrants associates to 1.15% increase in the density of formal establishments and 9.11% for the informal ones. This reveals that capital (in the form of new establishments), adjusts to the immigration-induced labor supply shock. Furthermore, this positive relationship is driven almost exclusively by an increase in the number of small establishments with less than 10 employees. The results from the industry analysis, confirm that immigration has the strongest impact on establishments in relatively mobile industries. In contrast, immigration has an insignificant impact on the number of establishments in high-skill services (less mobile)[2].

It is also found strong spillover effects in the expansion of informal firms, being 5 times higher compared to the formal sector where the spatial indirect effects are null. Thus, there exist evidence that immigration has differential association with business formality and informality throughout the city. Also, it is found greater capacity of immigration to increase formal firms in zones near the CBD; While, in the west and south east for the informal ones.

This study is aimed to make three contributions: first, to my knowledge, studies of how developing cities approach migration where informality is large, are scarce. This type of analysis is even more scarce for the case of Venezuelan migration. Second, most studies, either for developed or developing countries, have focused on migration effects from the labor supply side, that is, taking as dependent variables the levels of employment or native unemployment, but little regarding how labor demand reacts to a migratory shock and the nature of this adjustment.

The reminder of this paper is organized as follows: In section 2, we present the literature review about the topic, whereas in section 3 we introduce our theoretical perspective based on a model of specialization of land. Section 4 presents the empirical setting and section 5 shows the descriptive data and the spatial exploratory analysis. Section 6, discusses the estimation results and section 7 the conclusions.

2 Literature review

Migration literature has mainly focused on immigration effects in advanced economies labor markets where formal sector is large. Few works analyze immigration in developing countries and how natives adjust to foreigners when most of the economy is informal (Viseth, 2020). Empirical studies often work on the United States and Europe cases (Altonji and Card, (1989; 1991); Schoeni, 1997; Card, (2001, 2007); and Card and Lewis, 2007) for the US labor market; (Pischke and Velling (1997) and Glitz (2012) for Germany; and Winter-Ebmer and Zweimüller (1996)) for Austria. Around 80% of business and labor in developed economies are formal, while in developing countries more than 70% of the employment is provided by the informal sector (World Bank, 2020).

Considering the case of developed countries, the canonical framework of analysis predicts the principal pushing factor for migration is job searching in host economies. This condition entails the increase of labor supply that under a context of a high substitution of labor, generates competition between the native and immigrants that can lead to reductions in wages and global employment. This phenomenon is more likely to occur among the less educated workers in low-skill occupations because these jobs are less specialized and routined (Ottaviano and Peri, 2012; Peri and Sparber, 2009; Amuedo et al, 2011).

In contrast, more specialized jobs require high skills and complementarities among immigrants and natives arise more likely to increase productivity and wages. However, when testing this theoretical hypothesis, the vast majority of empirical results for developed countries find that migration has not been associated with changes in wages or global employment (Friedberg (1995 and 2001); Grogger,2008; Kerr, 2011; Longhi et al., 2005; Longhi, 2010). Taking these results as given, this leads to the question about what mechanisms operate so that in real life, migration has little effect on employment and wages.

Edo (2019) supplies an extensive review of the literature to explain that absorption capacity occurs when, either, capital accommodation is possible or firms make technological adjustments.

[2]Mobile refers to the definition of as those economic activities able to relocate or expand when facing a labour shock. An example of these are low skilled occupations such as farmworkers, construction workers, and grocery clerks. Less mobile activities are mostly related to high skilled activities such as managerial, professional, technical and associated professional occupations.

Both events that can be materialized in developed economies. In particular, one mechanism from the labor demand side consist in that firms individually adjust to changes in the skill composition (induced through immigration) by expanding those industries unskilled labor intensive (Leamer and Levinsohn, 1995, Bhagwati and Srinivasan, 1983). As a result, firm profits rise, further increasing the demand for low-wage workers and driving up wages and employment.

The second mechanism is that technology adjusts in a way that the relatively more abundant type of labor is used more intensively (Lewis (2011, 2013); Peri and Sparber, 2009). Some authors (e.g Unel,2018); Fairlie and Meyer (2003); Evans and Jovanovic, 1989; Cagetti and De Nardi, 2006; Sarker and Unel, 2017) show that savings in labor costs derived from the labor supply shocks allows host economies to reinvest in other production processes that require improvements (e.g changes in equipment and technology used in manufacturing, improvements to the tools, techniques, software solutions used to assist in the supply chain and delivery system).

With improvements, new technology also expands possibilities for hiring more labor, reducing negative impacts of immigration on wages. Constant and Zimmerman (2006), Azoulay et al (2021) find that business ownership could be higher for immigrants than natives as they are “pushed” into entrepreneurship by survival due to poor labor markets opportunities. Olney (2013), Ottaviano and Peri (2012) consider that capital can adjust very quickly in the short term to migratory shocks in such a way that firms expand their capacity to absorb excess supply of labor and reduce the negative impact of immigration on wages.

They make explicit how in the developed world it is possible to operate the intensive margin, that is to say, how firms at individual level respond to changes in the local labor supply, rearranging their technologies and the skill intensity to accommodate the influx of low-skilled workers within industries (Lewis (2003 and 2005); Gonzalez and Ortega, 2008; Dustmann and Glitz, 2011).

In the case of developing countries, those mechanisms failed. Being informal brings implications for migrants related to the productivity gap with respect to developed countries. Informal sector is generally composed by small firms with low productive capacity (Nicolini and Roig, 2019). They cannot adopt new technologies to raise productivity and accommodate in the long term a shock of migrant labor at fair wages. The quality of employment is also affected, because informal workers are not covered by social security, either contracts with legal protection without job stability or permanence (Combarnous et al, 2019).

Reviewing the literature for developing economies, there is a first line of authors whose findings are similar to those of developed countries framework; not significant effects of migration on employment and wages neither in the formal and the informal sector in their analysis cases (Santamaría, 2020; Bahar et al, 2020; Tribín and Uribe et al. 2020; Morales-Zurita et al., 2020; Tribet al, 2020; Howard, 2019). There is another research line that finds small negative effects of migration on employment and wages given the size of informality in these countries (Lebow, 2020; Verme and Schuettler 2021; Altindag et al. 2020; Aksu et al., 2018; Ceritoglu et al., 2017; Tumen, 2016; Del Carpio and Wagner, 2015). The oversupply of labor from migration generates unemployment and reduces wages in a highly informal economy, which is usually linked to low-skill occupations of low productivity, where substitutability is larger.

Some authors (Delgado et al, 2022; Kleemans et al, 2018) find that a migration shock generates differential effects between formal and informal sector. Higher levels of unemployment are produced in the formal sector because the minimum wage is rigid, but flexible in the informal. This leaves unemployment in the informal sector unaffected with flexible wages falling. It is highly likely that a low skill native worker becomes unemployed when is replaced by a immigrant, under a minimum wages context in the formal sector. Unemployment is mitigated by the transit of workers from the formal to the informal sector, where, everybody is welcome. Substitution among workers is large infinitely, wages adjust in such way that even under low productivity everybody accommodates inside informality.

Other alternative absorption mechanisms found in the literature are more focused on showing how clustering race or ethnicity has impacts on the location of employment and business niches in a city. However, these applications are still for developed countries in North American cities. Pioneering authors like Ellis et al, (2004, 2007) developed the concept of niches. The migration effects are reinterpreted in terms of how the expansion of immigrants cluster business and influences the formation of economic niches. What happens to immigrants is similar to what happens due to race effects, black professionals are often peripheralized in workplaces dominated by whites (Anderson, 2001; Bell and Nkomo, 2001; Waldinger and Der-Martirosyan, 2001; Wright and Ellis, 2001; Wilson, 2003; Wang, 2004). The network theories with a spatial perspective emphasizes on home-work geographies for local labor market outcomes (e.g., Stoll and Raphael, 2000; Logan, Alba, and Zhang, 2002; Parks, 2004; Wang, 2006) and show that living in immigrant neighborhoods rises the probability of working in a niche-sector job, so those neighborhoods are associated with a type of economic activities in space.

Another mechanism is pointed by Olney (2013) from the demand side. It is the expansion of the number of firms/establishments or changing the existing ones, which could explain the absorption of labor surplus. Facing cheaper labor, firms are motivated to produce at a lower cost and natives to run new businesses. According to Olney, since new establishments lead to an increase in the capital stock, the labor to capital ratio and thus wages are relatively unaffected.

In this wave are Constant and Zimmerman (2006) who compares the performance of the immigrants with those of the West native Germans from a socioeconomic Panel of year 2000. Immigrants or natives will choose to be self-employed if the expected earnings from self-employment exceed the expected earnings from other types of employment. Altindag et al (2020) also studied the entrepreneurship creation analyzing the effects of the forced migration of Syrians to Turkey on firm creation during the period 2004-2016. These authors find that one-percentage-point increase in the share of refugees (as a percentage of population) leads to 1.5 percentage-point increase in the number of Turkish firms. The Syrian refugees replaced native workers in the informal labor market and reduced labor costs for firms.

Another issue to be reviewed in the literature is the heterogeneity of the informal sector. Two important facts are mentioned by Ulyssea (2018, 2020). First, formality and informality coexist even inside “the same industries producing similar products ” (Ulyssea, 2020, page 527). Both sectors interact in the production process, either as the inputs purchases or products sales, through production chains, they are not exclusive, rather there is a dynamic between them. Also, there could be transitions; an informal firm could become formal after a shock that affects positively profitability. Either a formal firm could become informal to safeguard a negative macroeconomic shock.

Second, heterogeneous firms sort into the formal or informal sector. According to Ulyssea (2018), three types of informal firms can be classified: the potentially productive ones that need to start informally to grow. Parasite firms that already have the productivity to become formal but choose to remain informal and earn higher profits to avoid costs with taxes and regulations. And the informal firms with low subsistence productivity that do not develop their potential. In the experience for Brazil, the author cites that only 9.3% of firms are in the first group, 41% in the second and 48% in the third. The great weight of the latter is what justifies modeling the associating between the informal sector with low productivity. To complete the author argument, the larger the firm, the less likely it is to become informal because of the difficulty to get out of the spectrum of the authorities and more dependent on the financial system.

The mentioned heterogeneity of the informal sector is reflected on the territory. Economic activities such as human choices express their evidence on the footprint occupation. Hernandez et al (2011), show that the Central Business Districts of developing countries are characterized by a high mix of formal and informal activities. As the distance from the CBD increases the rate of informality grows. These small potential and the parasite firms organize into space according to their budgets and productivities. The largest ones are located into the Central Business District (CBD), since they can pay higher land rents. The small (formal and informal) will be located at a middle distance. The informal ones with low productivity at a greater distance from the CBD. It is not about an exact categorization of the location, but there can exist multiple combinations in which the location as well as the degree of formality and size flow between the periphery and the city center [3]. Thus, when migrants arrive into host developing cities what found is theese kinds of spatial structure of the economic tissue.

Not only firms but also workers sort spatially. Posada et al (2019); Muñiz et al (2020) assume that formal workers have higher commuting frequencies, although they have higher budgets to live close to the city center. Informal workers commute less frequently to the CBD than formal workers. Their low informal wages make their valuation of being close to the CBD lower in such a way that informal workers end up living at the periphery. This could generate patterns of segregation in which workers (natives and immigrants) and informal firms are confined to the peripheries.

The economic spatial structure shapes the location of immigrants and native jobs. In that order, there is also a spatial distribution of immigrants between formal and informal firms. The entry of these immigrants generates downward pressure on informal wages, undoubtedly increasing the number of informal businesses. There is also an effect on formal firms: the entry of (informal) immigrants, hired directly by formal firms, due to lower labor costs, can expand or create new establishments. Alternativelly, the case of informal immigrants who are hired by informal units, increases in size, stop being informal and become formal. In other words, an impact of formal and informal business dynamics (business creation) is identified.

However, the existence of spatial segregation does not prevent interdependencies between formal and informal firms and workers in space. The proliferation of informal businesses in one area can encourage the creation of firms in the surrounding areas. If the expansion of firms occurs in part due to migration inflows, this relationship can also be extended through the space forming a cluster or areas where migrants play a key role in entrepreneurship.

3 Theoretical setting

Here, we introduced a simple theoretical setting based on land specialization model proposed by García-López et al (2020) and Flores-Fillol et al (2016). These authors propose a structure based on the idea that CBD settles in the origin of a real line $X \in (-\infty, \infty)$ where clusters high-value services locate. The CBD is a point of attraction for which each type of agent in the city competes for settling close.

Let us define any distance $x \in X$ from the CBD, with $x > 0$. Following Flores-Fillol et al (2016), here it is modelled three types of agents competing for land near the CBD: (1) a continuum of identical formal firms with density $n_f(x) \geq 0$ at $x \in X$; (2) a continuum of identical informal firms with density $n_i(x) \geq 0$ at $x \in X$; (3) a continuum of consumers-workers with density $n_c(x) \geq 0$ at $x \in X$. The subscripts “f,' ' “i,' ', “c' ' denote formal, informal firms and consumers, respectively. As in the classical models, an absentee landlord is assumed. The total land is finite, occupied

[3]In addition, if the productivity of the formal sector increases, it should be extended through a greater demand for goods, across space towards the periphery. If productivity in the periphery improves, it should be extended in the opposite direction through lower input costs and quality for the formal ones. This is to say, the existence of spillover effects.

by the three agents at $x \in X$ and normalized to one (here we follow Cavailhés et al (2004) and Flores-Fillol et al (2016)):

$$n_f(x)S_f(x) + n_i(x)S_i(x) + n_c(x)S_c(x) = 1 \tag{1}$$

where $S_f(x)$ is the size of land plot occupied by formal firms in the city, $S_i(x)$ is the size of the land occupied by the informal and $S_c(x)$, are the size of the consumer-workers plot respectively.

3.1 Agents behavior

In this framework, formal firms produce a good Y_f using inputs supplied by the informal sector $Y_i(x)$ at each location $x \in X$ with land $S_f(x)$. Informal agents also produce a good Y_i using inputs supplied by the formal sector Y_f at each location $x \in X$ and land $S_i(x)$. Both sectors, formal and informal, are connected through input demands. Their production function is modeled by a Cobb-Douglas with constant returns to scale. The action of buying from formal or informal firms implies the existence of transportation costs. For the sake of simplicity, we consider that transportation costs are partially assumed by both firms type and are proportional to their distance from the CBD.

Formal and informal firms sell (settled at $x \in X$) their goods at a price p_f and p_i respectively. Revenues earned by formal and informal firms are discounted by their transportation costs t equal $\frac{p_f}{tx}$ $Y_f(x)$ and $\frac{p_i}{tx}$ $Y_i(x)$ respectively. $R_f(x)$ and $R_i(x)$ are the rents paid by the formal and informal firms at location x. Formal firms choose $S_f(x)$ and $Y_i(x)$ to maximize profits. Informal firms choose $S_i(x)$, $Y_f(x)$, to maximize their profit. Each of the maximization problem faced by firms type and their first order conditions are shown in table 1, equations 6 to 9.

Consumers choose the optimal level of two composite goods; the formal good $Y_f(x)$ and the informal one $Y_i(x)$ as well as land $S_i(x)$ to maximize their utility. Consumers obtain their income as an endowment w. Commuting to the CBD implies incurring in transport costs t (with $t > 0$). The first order condition yields equations 2 to 5 in Table 1.

Table 1. Maximization problems of agents

Consumers Behaviour(1)	Formal firms(2)	Informal firms(3)
$Max\ U_c = Y_f(x)^\alpha Y_i(x)^\beta S_c(x)^{1-\alpha-\beta}$ s.t w=$p_f Y_f(x) + p_i Y_i(x) + R_c(x)S_c(x) + tx$ with $\alpha, \beta \in [0,1]$	$Max\ \pi_f = \frac{p_f}{tx} Y_i(x)^\gamma S_f(x)^{1-\gamma} - p_i Y_i(x) - R_f(x)S_f(x)$ with $0 < \gamma < 1$	$Max\ \pi_i = \frac{p_i}{tx} Y_f(x)^\alpha S_i(x)^{1-\alpha} - p_f Y_f(x) - R_i(x)S_i(x)$ with $0 < \alpha < 1$
First Order Conditions	First Order Conditions	First Order Conditions
$S_c^*(x) = (w - tx)[\frac{1-\alpha-\beta}{R_c(x)}]$ **(2)** $Y_f^*(x) = \frac{\alpha(w-tx)}{p_f}$ **(3)** $Y_i^*(x) = \frac{\beta(w-tx)}{p_i}$ **(4)** $R_c^*(x) = [\frac{(w-\alpha tx)}{p_f^\alpha V}]^{\frac{1}{1-\alpha}}$ **(5)**	$S_f^*(x)\beta\frac{(w-tx)}{p_i}[\frac{p_f(1-\gamma)}{txR_f(x)}]^{\frac{1}{\gamma}}$ **(6)** $R_f^*(x) = (1-\gamma)\frac{p_f}{tx}^{\frac{1}{(1-\gamma)}} \frac{\gamma}{w}^{\frac{\gamma}{(1-\gamma)}}$ **(7)**	$S_i^*(x) = \alpha\frac{(w-tx)}{p_f}[\frac{p_i(1-\alpha)}{txR_i(x)}]^{1-\alpha}$ **(8)** $R_i^*(x) = (1-\alpha)\frac{p_i}{tx}^{\frac{1}{(1-\alpha)}} \frac{\alpha}{p_f}^{\frac{\alpha}{(1-\alpha)}}$ **(9)**

Own elaboration.

Equation (6) reveals that an increase in transport costs reduces the consumption of land. An increase in the marginal revenue of the formal firms $\frac{p_f}{tx}$ encourages the firm to occupy more land. Finally, the land plot input decreases with the rent cost $R_f(x)$. The same applies for the informal firms analysis in equation (8). Competition for land is assumed to extract all profits (zero-profit condition) which allows us to calculate the maximum level of rent to be paid by formal or informal firms respectively. The increase in labor costs, w , reduce the ability of formal firms to pay high rents close to the CBD. The rent function (Eq 7) increases with marginal revenue $\frac{p_f}{tx}$, and decreases with respect to the distance to CBD.

For the informal firms, the first order condition appears in column 3. Equation 6 tells us that increases in transport costs, reduces consumption for land, $S_i^*(x)$. The rental cost $R_i^*(x)$, exerts a negative pressure on the individual's land consumption.

3.2 Equilibrium full specialization of land

Following Fujita and Thisse (2002), we assume that agents compete for land as in an auction mechanism. At the equilibrium, the maximum rent at each location is the one offered by the highest bidder. As a consequence, the rent curve at $x \in X$ is an envelop $R^*(x) = Max\{R^*_f, R^*_i, R^*_c\}$. For this reason, land is specialized and given to the highest bidder at any point x. After the bidding process ends no land is vacant. Thus, looking at equation 1, it is possible to check that there is land specialization. Thus, for the only area occupied by the formal firms the density can be written as $n_f(x) = \frac{1}{S^*_f(x)}$. For the informal firms $n_i(x) = \frac{1}{S^*_i(x)}$. The only residential area holds as $n_c(x) = \frac{1}{S^*_c(x)}$

Taking equations 2 and 6, and using the fact that $n_c(x) = \frac{1}{S^*_c(x)}$ and $n_{f(x)} = \frac{1}{S^*_f(x)}$ we obtain :

$$n_f(x) = \frac{p_f(1-\alpha-\beta)R_f(x)}{\alpha\frac{p_f}{tx}(1-\gamma)R_c(x)} n_c(x) \qquad (10)$$

The same applies to the informal sector. Taking into account equation (8) we obtain the following:

$$n_f(x) = \frac{(1-\alpha-\beta)}{R_c(x)\alpha} \frac{p_f R_i(x)^{\frac{1}{\alpha}}}{\frac{p_i}{tx}^{\frac{1}{\alpha}}(1-\alpha)^{\alpha}} n_c(x) \qquad (11)$$

Here, we express the formal and informal densities in terms of the consumers densities. Given the interrelationships that exist in the model between the formal (informal) sector and consumers through the consumption of formal (informal) goods, we obtain an expression that allows us to illustrate how the increase in the density of consumers contributes to an increase in the density of formal (informal) firms. This is because the greater the demand, the greater the production for the informal sector and the greater the density of informal firms in their specialization and fixed land plot[4].

3.3 Disaggregating by consumer type

The Flores-Fillol et al (2016) model gives us a framework to show how, through the demand side of firms and the increase in the set of consumers, it encourages an increases in the density of both formal and informal firms under a scenario of competition for land. The framework allows to interpret at this point, what would have happen if we could disaggregate $n_c(x)$ into two types of consumers; native or immigrant consumers. The framework answer would be to capture a migratory shock. For instance, an increase in the density of immigrant consumers, who at the same time are workers (either the density of native consumers), will result in a positive effect on the density of formal and informal firms.

[4]Take equation 2 and replaces $w - tx$ by the expression $Y^*_f(x)\frac{p_f}{\alpha}$. This expression comes from Eq3 obtaining:

$$S^*_c(x) = Y_f(x)(\frac{p_f(1-\alpha-\beta)}{\alpha R_c(x))}$$
$$S^*_c(x) = Y_i(x)^{\gamma} S_f(x)^{(1-\gamma)}\frac{p_f(1-\alpha-\beta)}{\alpha R_c(x))}$$
$$S^*_c(x) = \frac{Y_i(x)^{\gamma}}{S_f(x)^{\gamma}} S_f(x)\frac{p_f(1-\alpha-\beta)}{\alpha R_c(x))}$$

And now we now that $n_c(x) = \frac{1}{S^*_c(x)} = \frac{S_f(x)^{\gamma}}{Y_i(x)^{\gamma}} n_f(x)\frac{\alpha R_c(x)}{p_f(1-\alpha-\beta)}$. Finally, replace $S_f(x)$ and obtains Eq 11

4 Data

4.1 Venezuelan exodus context

We have selected the migration case of Venezuelan, because it is the most massive and representative case of migration from and to a developing country like Colombia. It is the country that receives most migrants from Venezuela with 32%, followed by Peru (17.39%) and Chile, (9.67%)[5].

It is important to understand the features and magnitude of this migration. The origin of the Venezuelan exodus has been widely documented by several authors (e.g Freitez (2011); Lacruz (2006); Crasto (2017)) having its start in the year 2000, when Hugo Chavez came to power. With the nationalization of the oil company PDVSA, the threat of expropriation of the private property unleashed the first migratory wave of early 2000, with the upper class; a small number of industrialists and politicians who felt threatened by the Hugo Chavez govern. The second wave was until 2005 and took on a more worrying hue, when the Chavez government fired 18.000 employees of the official oil company PDVSA and took control of it (Freitez (2011)).

Thus, many skilled workers in the oil field emigrated given the repressive measures of Chavez such as disqualify them for their whole life from working with the public sector. This situation was exhausted by the regime of Nicolas Maduro, who assumed power in 2014 and prolonged the Chavist measures in a more repressive way. An unemployment crisis was unleashed, exacerbating insecurity in the streets, food and medicine crises for ordinary citizens. This is one of the reasons for the third wave of migration in 2015 to Colombia. According to the official figures of Migration Colombia Office, arrived into the national territory about 329.478 Venezuelan citizens (Migracion Colombia (2017)). From 2014-2022 Migration Colombia Office registered about 2.3 million Venezuelans have entered the country officially. However, it is estimated that in an irregular form there could be around one million more. The 42% of this population is concentrated in the 5 largest cities of the country: Bogota, Medellin, Cali, Barranquilla and Cartagena followed by border cities such as Cucuta, Maicao, Riohacha and Bucaramanga.

4.2 Inmigrants and their Distributional location in Bogota

We define immigrants as those who were born in Venezuela and reported the year when they entered the country within period 2012-2022. Data for immigrants between 2012 and 2018 are built from the Census of Population and Housing 2018 made by the Statistical Department of Colombia (DANE in Spanish). It contains a question about the exact entry year, hence, we can track immigrants backwards. From 2019 up to 2022, data have been complemented with information from the Register of Venezuelan Migrants (RUMV in Spanish).

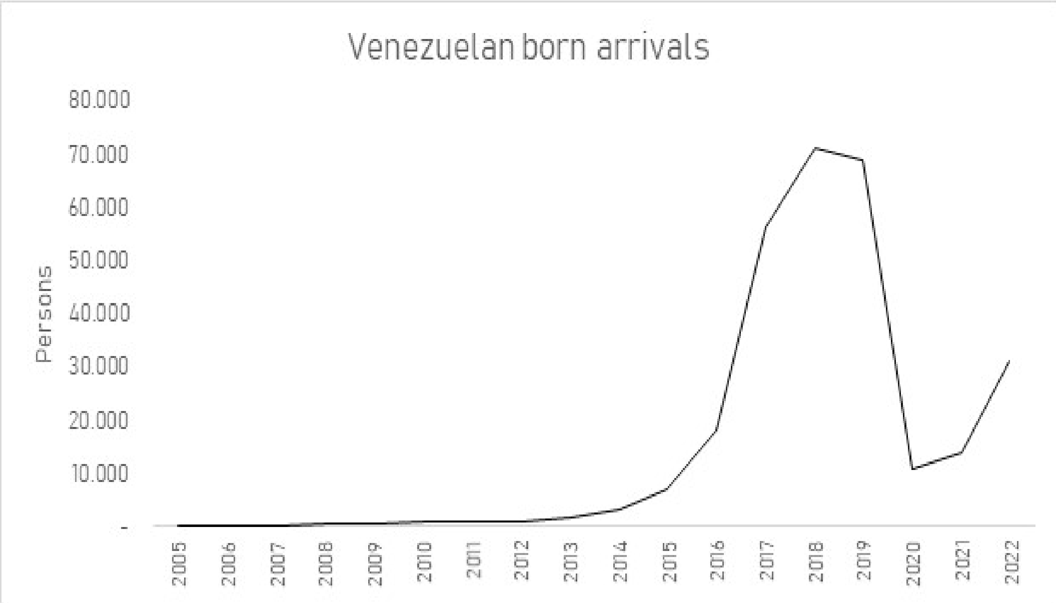
The RUMV is a virtual registry the government makes available online to Venezuelans in order to obtain socioeconomic information for public policy design. Registering on RUMV is also an incentive for immigrants to obtain a Special Stay Permit (PEP in Spanish) to reside and work in Colombia. It contains information of regular and irregular immigrants that decided to be registered and stay in Colombia before 31 of january of 2021. In additon, the Identifying Potential Beneficiaries of Social Programs (SISBEN in Spanish) is also used as complementary data. All individual information is aggregated at the level of 1.171 neigborhoods in Bogotá.

In cumulative terms, between 2005-2022 are identified spatially 291.011 immigrants. Figure 1 shows the flows of migratory arrivals across 2005 and 2022 in Bogota. According to Figure 1, the migratory peak occurred between 2018 and 2019 and gradually the arrival flows decreased towards the year of pandemic 2020, date in which the borders were closed due to the health emergency.

[5]Figures from the Migration Office Report 2022

Progressively immigration increased slowly until 2022 but not recovers the achieved flows before pandemic year.

Figure 1. Number of Venezuelan migrants by birthplace and arrival year to Bogota



Source: Information before 2018 was built on Census of Population and Housing-DANE and from 2019 up to now from administrative registers.

The official information reported by the Statistical Department of Colombia-DANE (see table 2) estimates the stock but not the flow of migrants, based on a Labor Market Survey (the Great Integrated Household Survey). A larger average stock number of migrants are exhibited in the survey until 2022 compared to the ones calculated through the administrative records above. That is to say, the stock number of Venezuelans estimated from the Great Integrated Household Survey-GEIH counts for around 323.205 migrants from last period 2020-2022 while from mixed data sources it is found 291.011.

This happens because the GEIH is a survey that expands the population with statistical criteria that does not necessarily corresponds to the total number of migrants, while the census and administrative records are data that is effectively visible to the authorities. The foregoing means that of the 323.205 stock cumulated migrants estimated from GEIH, it is only possible to spatialize 291.011 cumulated persons across 2012-2018, recovered from the Population Census 2018 and from administrative records. Table 2, also shows the highest rates of unemployment and informality that face Venezuelan migrants, compared to the native citizens.

The next is to get closed about where migrants located to live in the city. At spatial level, map 1a below, shows the spatial distribution of native population at neigborhood level as a reference point for immigrants when arriving. The highest native population concentrates in the west-south periphery with a density that reaches 399 thousand people per square kilometer. In the internal Northeast areas, the densities are slightly lower between 11 thousand and 45 thousand people per square kilometer. This as a result of the city’s own socio-spatial segregation since the 50s, in which the poor population lives in overcrowding while higher income housings live in lower densities.

Table 2. Descriptives statistics of Migrants and Natives population in Bogotá

Period	2013-2016		2017-2019		2020-2022	
Migrants						
Type	Mean	S.D	Mean	S.D	Mean	S.D
Population	27.769	11.894	243.990	56.235	323.205	43.997
Employment	18.435	5.993	125.181	67.390	160.298	19.352
Economic Active Population	22.507	8.656	155.075	85.605	194.249	31.933
Workforce	24.472	9.503	194.719	19.068	222.106	79.127
Informal workers	13.035	4.248	103.776	65.378	142.365	17.752
Occupation rate	77%	6%	66%	7%	76%	16%
Unemployment rate	17%	5%	19%	4%	17%	4%
Informality rate	71%	5%	79%	12%	89%	1%
Natives						
Population	7.227.200	39	7.347.000	49	7.690.000	122
Employment	4.190.743	7.783	4.165.244	47.328	3.687.612	10.915
Unemployment	409.598	16.581	485.280	1.408	658.485	16.246
Economic Active Population	4.600.431	15.604	4.650.524	47.100	4.346.097	92.539
Workforce	6.424.188	98.682	6.717.403	99.856	6.567.601	30.708
Informal workers	2.328.785	48.024	2.022.144	28.624	1.754.002	69.589
Occupation rate	65%	1%	62%	0%	56%	3%
Unemployment rate	9%	0%	10%	0%	15%	4%
Informality rate	43%	1%	42%	1%	39%	4%

Source: GEIH-DANE

Map 1a also describes the main central areas of residential attraction that population wishes to locate as strategic to live or work. These correspond to the colored polygons named Santa Bárbara, Sagrado Corazón, San Diego, Chapinero, Chico, Modelia and Salitre. These same areas constitute poles of reference for the location of migrants

Maps 1b-1d, show the distribution for immigrants between the period 2012 and 2014, who tended to settle in the north-eastern part of the city, areas corresponding to high socioeconomic status levels. This coincides with the arrival of the first wave of high-income migrants. Since 2015 the location tends to be more peripheral towards the western outskirts of the city. After 2018, concentration of migrants is much greater towards the south-western periphery due to low-income immigrant population.

The immigrant population with a higher educational level is also located closer to the CBD. About 47.000 migrants with higher education levels (university or postgraduate), chose their location no less than 4 kilometers from the CBD. Likewise, those of medium education that correspond to technical education. For their part, migrants with less than an educational level, despite being located even close to the centers of population attraction, tend to concentrate more than 9 kilometers from the CBD.

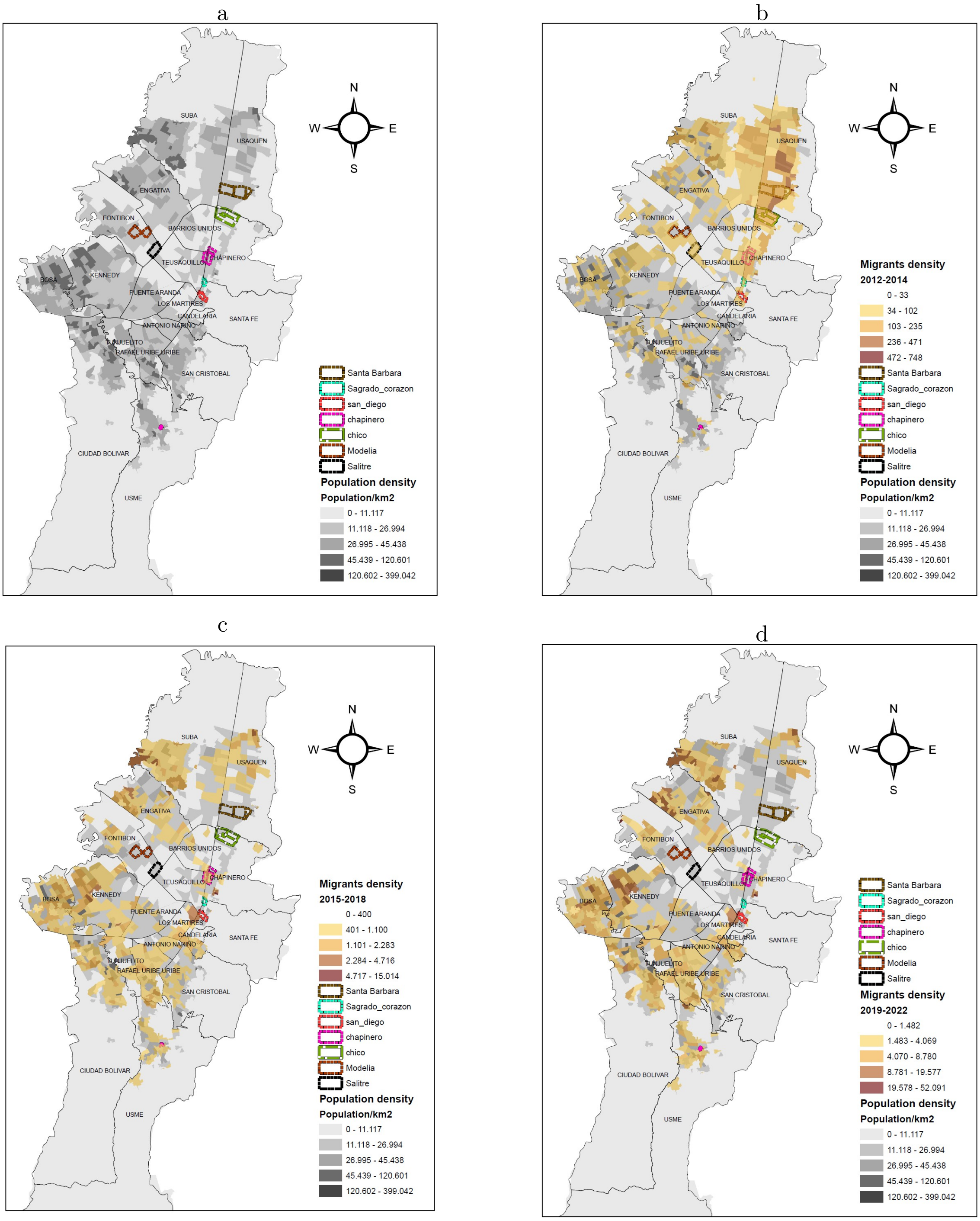
4.3 Distributional location of firms in Bogota

In this document, we introduce the definition of formality by the International Labor Organization-ILO as those firms which do pay social security (health and/or pensions) to their workers.

4.3.1 Administrative Registers- for firms and formal workers

The Integrated Form of the Payment Contributions (Planilla Unica Integrada-PILA in Spanish) is used as a source of information for building the labor formality definition and data is available for the period 2018-2022. This source contains information of formal firms contributors who affiliated

Map 1. Native and migrant population location along the city by sub periods 2012-2014, 2015-2018, 2019-2022 and neighborhoods.



Source: Before 2018 Census of Population and Housing-DANE and from 2019 up to now from administrative registers.

their employees with social security such as pensions, health, Occupational Risks, Family Compensation Funds, parafiscal payments. Also, it is possible to retrieve information on the existence of the firm taking as reference the question to the worker about “since what year have you worked in this company”. PILA currently has 170.630 companies, of which 164.065 can be geo-referenced and aggregated at neighborhood level.

For the second source of informacion we have the administrative records of the Bogota Chamber of Commerce available from 2012 to 2022 annually. Every time a company is formally created, it must be registered with the Chamber of Commerce to become commercially active in the market. Each year it renews its commercial registration.

This database contains the records of the companies, created, renewed or cancelled registration. It also contains information on the company’s physical address, number of branches, assets, sales, reported jobs, company size. Thus, it can be georeferenced at the block or neighborhood level. We used this database to track formal firms found in PILA from the missing period 2012-2017.

However, none of these sources capture informal firms. That is the reason why, in this study we calculate a “potential informality indicator” by using two censuses of economic units; one carried out in 2017 by the local mayor’s office and another developed by the National Statistical Department-DANE in 2021. Both censuses have information of the total number of firms but do not discriminate among formal/informal firms. Hence, the potential informality indicator is calculated as the difference between the total number of censal firms and those previously registered as formal firms on the administrative records. Also, Economic Census of 2017 include the question. “¿ In what month and year did this economic unit started operations?”, which allows companies to be located in the years in which they were created and their existence over time.We assume the same Census structure of 2017 for 2018. For year 2019-2022 we used Census of Economic Units of 2021.

Table 2. Descriptives statistics of Migrants and Natives population in Bogotá

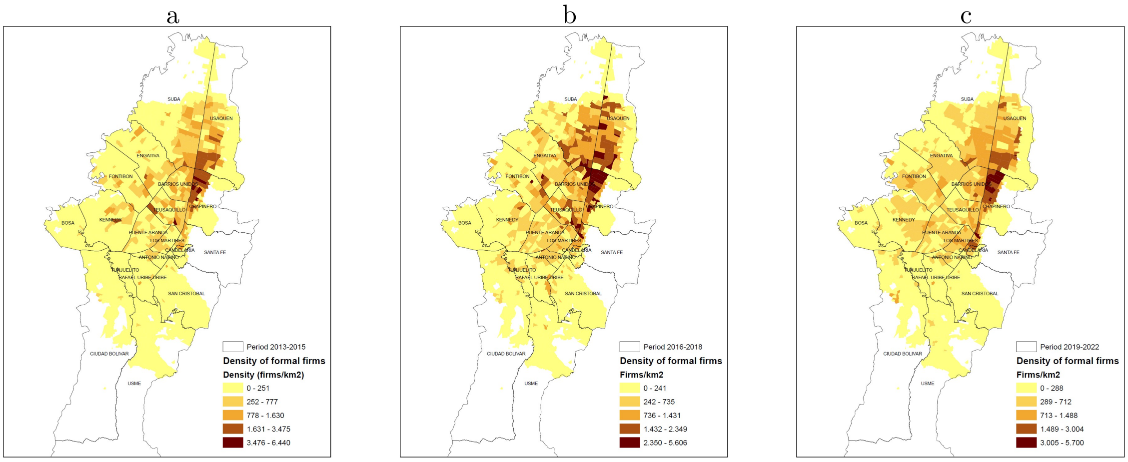
Type	2013-2015	2016-2018	2019-2022
Formal	113.503	144.471	153.575
Potentially Informal	233.931	349.996	357.914
Total	347.433	494.467	511.289
Potential Informal Rate	67%	71%	70%
Official Informality Rate by register			70%
Official Informality Rate by social security			75%

Source: Formal firms from Chamber of Commerce of Bogota and the Integrated Form of the Payment Contributions. Informal firms calculated from Economic Census 20187-2018. *Official statistics from National Department of Statistics-DANE based on the Multidimensional index of informality.

Table 3 shows the total number of firms at the city level, that have managed to be identified by crossing the different sources of information for time periods 2013-2015, 2016-2018 and 2019-2022. The potential informality indicator is between 67%-70% inter-period. This percentage coincides with the official authorities-DANE for both business informality (the indicator is 70%), and labor informality (the indicator is 75%). For the periods 2013-2015 and 2016-2018 there is not previous measurement of business informality at the official level.

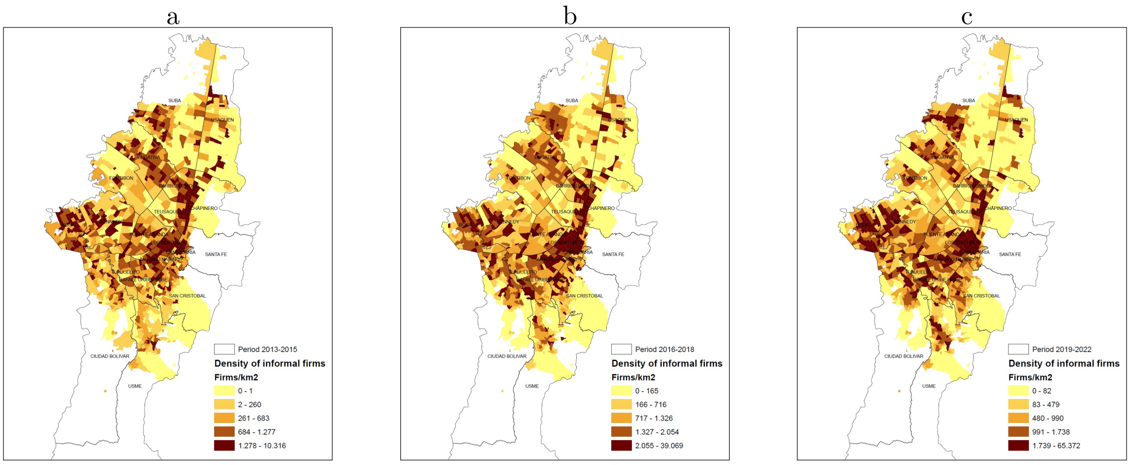
At spatial level, there are two patterns of location of the firms, the formal location pattern cluster to the north of the city (map set 2) and the informal ones cluster to the south.

Map 2. Formal firms location along the city and periods 2012-2014, 2015-2018, 2019-2022 and neighborhoods.



Source: Formal firms from Chamber of Commerce of Bogota and the Integrated Form of the Payment Contributions. Informal firms calculated from Economic Census 2017;2021

Map 3. Informal firms location along the city and periods 2012-2014, 2015-2018, 2019-2022 and neigborhoods

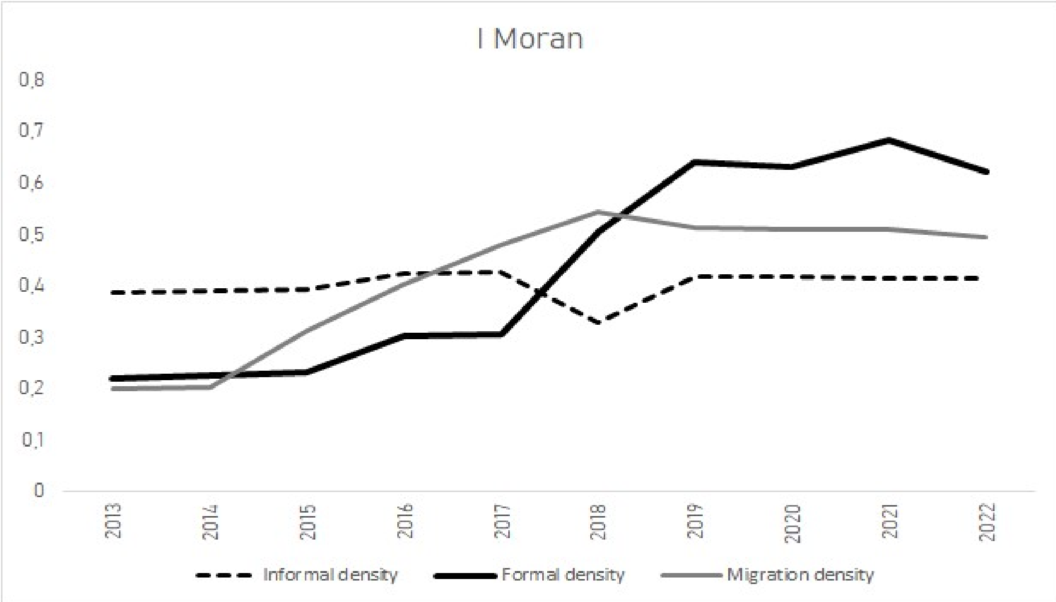


Source: Informal firms calculated from Economic Census 2017;2021.

4.4 Spatial Exploratory Data Analysis

We perform an exploratory spatial data analysis ESDA-type to verify whether the variables of interest, namely, densities of economic units and immigrants, exhibit spatial autoregressive connection patterns. We compute the local I Moran which is a measure of spatial autocorrelation.

Figure 2. I Moran of formal and informal firms and migrants (p-value<0.05)



Source: Own elaboration.

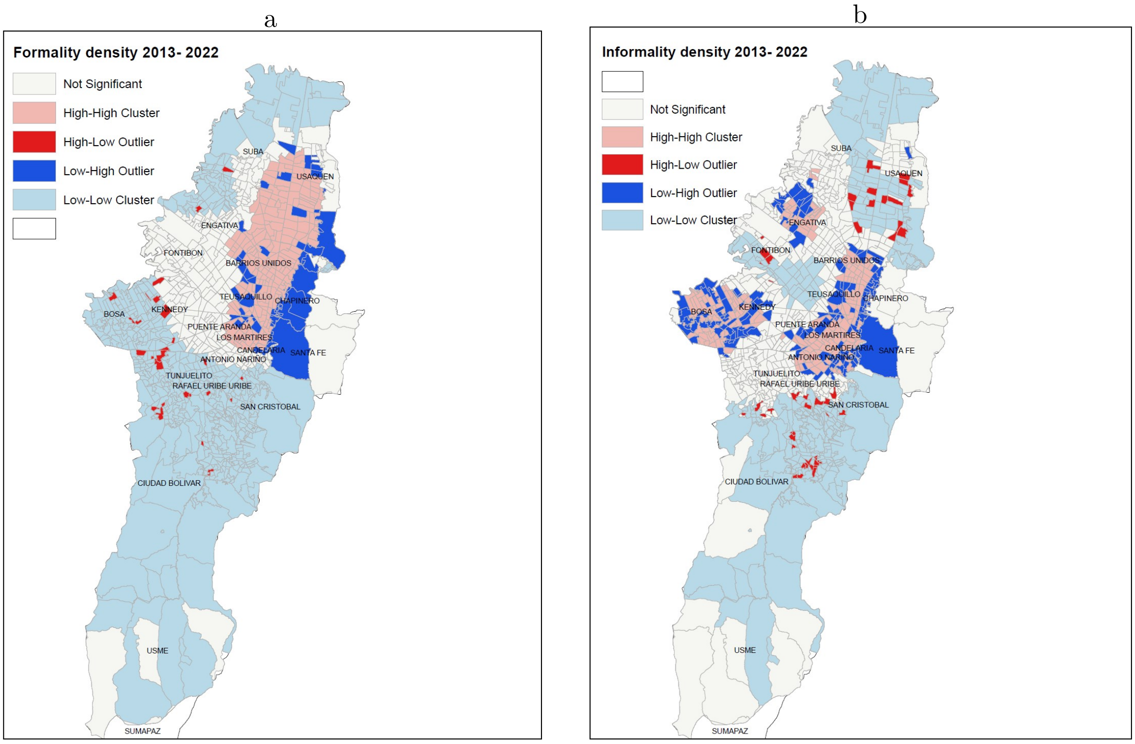
As one can deduce in figure 2, the univariate Local Moran Index for all the variables of interest is positive and significant over time. In the case of formal firms, the I-Moran is in a range of 0.2-0.7, denoting a high positive spatial autocorrelation, especially after 2017. There is a concentration of formal firms towards the central areas of the city as shown in Cluster Map 1a and which is denoted in pink color that corresponds to the High-High cluster areas. The spillover effects of the location of formal firms were reinforced in the territory, since 2017, with a slight decrease in the 2020 pandemic year. In the case of informal firms, the Moran I ranged around 0.4 throughout the study period, with a slight decrease in 2018. This means that informality is slightly more dispersed than formality as the segment of precarious informality is located almost in any place of subsistence.

Cluster Map 1b right side shows three important cluster areas with positive or High-High autocorrelation (areas with a high concentration of informal firms surrounded by neighboring areas with a high concentration of informal firms): the central zone close to the CBD of the historic center, towards the western zone and towards the south-western part of the city.

In the case of migrants per square kilometer (Figure 2), Moran's I oscillate in a range of 0.2 to 0.5, a progressive growth, which began in 2013 with a timid concentration in the north of the city and explains why a low Moran I of 0.2. A greater concentration of the migrant population appears in 2018 with Moran's I of 0.58 and then it tends to stabilize until now around 0.5. Moran's I is consistent with the gradual arrival of migrants from 2013 to its maximum peak in 2018 and then its subsequent decline.

The set of maps in Appendix Figure A1 shows how the clusters with positive High-High auto-correlation change over time, starting in 2013, moving from the north of the city to the north-west, then to the south-west and finally consolidating a cluster in the south of the city.

Cluster Map 1. Formal density 2013-2022 in neigborhoods.



Source: Own elaboration.

5 Empirical setting

We settle our empirical analysis on the case of Bogota city. We are interested in measuring the association between migration with the expansion in the number of firms. Here, we capture three key elements; the spatial structure of the city, the sectoral relationship among formal and informal sectors and its projection on the territory. We propose an extension of García-López et al (2020) model that incorporates the spatial distribution of firms in both the formal and informal sector. This monocentric structure influences when migrants or native firms make location decisions. The estimation strategy is based on the resulting equation 12, using logarithms as:

$$LnDEst_{ijt} = \alpha_{0i} + \alpha_{1i} LnDEst_{jt(h \neq i)} + \alpha_{2i} LnDImm_{jt} + \alpha_{3i} Lnx_{ij0} + \alpha_{4i} LnX_{jt} + \rho_i W LnDEst_{ijt} + \theta_{1i} W LnDEst_{jt(h \neq i)} + \theta_{2i} W LnX_{j0} + \theta_{3i} W LnX_{jt} + \delta_j + \theta_t + \epsilon_{ijt} \quad (12)$$

Where $LnDEst_{ijt}$ is the density of the establishments by km2 of the type i (i=Formal, Informal) in spatial unit j (here neighborhood) at time t. The key independent variable $LnDImm_{jt}$ is the number of immigrants by km2 coming from Venezuelan who arrive to spatial unit j in time t. Venezuelan immigrants represent 90% of the total immigrants in Colombia who reside more than one year for which is the predominant population of migrants in Colombia. x_{j0} is the distance from any spatial unit j to the CBD. The coefficient α_{3i} helps identify the degree of the attractiveness of the selected CBD and, hence, the possible monocentric spatial pattern of Bogota.

The identification of the CBD is one of the first issues that we must address to understand what is defined as the Central Business District-CBD and what is the logic of population location centers in Bogotá. In this particular case of study, we chose the CBD as one core of attraction by the location of the population, arguing that they are the first points of reference once an immigrant arrives, being the search for housing their first location decision. In appendix Table 1A, we present the estimates to choose the CBD of Bogota among several candidates. Following the procedure of García et al (2019) it is found that a place located at the north of the city called Unicentro was obtained as the CBD of Bogotá as the greatest population attraction, both for natives and for immigrants.

The vector $LnDEst_{jt(h \neq i)}$ represents the intersectoral relationship among sectors, specifically, the density of all other of establishmentâs $h \neq i$ in the same spatial unit j. Meanwhile, δ_j is a vector of neighborhood fixed effects while θ_t controls for time effects. The vector X_{jt} refers to other explanatory variables that are likely to explain the increase in the density of establishments in spatial unit the city that are not due to factors different from migration like the concentration of consumers (proxied by population density). Finally, ϵ_{jit} are the disturbance terms. The spatial weight matrix W is computed as a first order continuity matrix and all the terms of the equation that interact with W, correspond to the spatial lags of the variables.

Finally, we need to deal for endegeneity, as it is very probable that immigrants locate precisely in areas of greater density of firms and at the same time, firms expand in number due to the effect of the arrival of immigrants. To address this double causality we implement two strategies. First, we lag all the independent variables one period. That is, including a lag “Log Migration density(t-1)” would say that an increase in formal firms today does not influence the greater attraction of migrants yesterday. The same applies for the others explanatory variables.

Second, we introduce a Bartik-shift-shared instrument proposed by Altonji and Card (1991) and Card (2001), based on past settlementof immingrants. As well known, the first term of a Bartik instrument type is the share of Venezuelans in every neighbourhood in the city in very past year. In our case we will chose 2005, as we count on the economic census at spatial level of 2005. The first term is multiplied by yearly migrants normalized by the working-age population. We will choose the working age population of 2018, as this year was the highest peak of migration period.

The validity of the past settlement instrument relies on the fact that new arrivals to neighbourhood j are expected to be attracted by the network effects in that location, while current economic business creation in j are unlikely to be systematically related to lagged immigration shares (if those shares are lagged sufficiently). If this holds, then the instrument is valid because the immigrant location is related to new arrivals (relevance) but is not related to current economic conditions (exogeneity) since the economic trends are not highly serially correlated enough when we move away in time. The IV regression specification includes two types of instruments, a first step regression non spatial instrument and spatial instrument with a specification SDM type.

6 Results

In this chapter, we aim to estimate the correlation between immigrant inflows and expansion in number of firms (informal/formal). Following the strategy described in LeSage and Pace (2009) and Elhorst (2010), we start from the Spatial Durbin Model-SDM type of equation 12, as a general specification and test for alternatives. It incorporates the spatial lag of the dependent and independent variables capturing for local and global spatial effects of Venezuelan migration on the proliferation of formal and informal firms. The goal is testing the best model selection among different spatial specifications type such as Autoregressive Model-SAR, Spatial Error Model-SEM and Spatial Autocorrelation Model-SAC. After, checking whether the interdependence components are significant or not. Lastly, the SAC and SDM are non-nested, information criteria AIC-BIC can be used to choose the best model.

Table 4. Estimation of the effects of Venezuelan migration on the formal business proliferation

	Depvar: Log Formal density			
Indepvar :Log Migration density (t-1)	(SDM)	(SAR)	(SEM)	(SAC)
Local effects	0.115***	0.136***	0.157***	0.0712***
	(0.00924)	(0.00891)	(0.00963)	(0.00832)
Wx: Global effects	-0.124***			
	(0.0666)			
Direct effects	0.8097***	0.8754***		0.7654***
	(0.01086)	(0.0107)		(0.01083)
Indirect effects	0.451	0.2667*		0.1207
	(0.159)	(0.04534)		(-3.964)
Total effects	1.2607*	1.1421**		0.886**
	(0.0359303	(0.0525443)		(0.0878932)
Spatial rho	0.499***	0.299***		0.1992***
	(0.000112)	(0.000111)		(0.000364)
lambda (W*error term)			0.0957***	0.00157***
			(0.0000358)	(0.000101)
Variance sigma2_e	4.966***	20.60***	36.13***	43.32***
	(0.0650)	(0.0406)	(0.866)	(0.567)
Time effects	Yes	Yes	Yes	Yes
All controls+	Yes	Yes	Yes	Yes
N	11,710	11,710	11,710	11,710
R-sq	0.26	0.263	0.28	0.43
adj. R-sq	0.22	0.23	0.18	0.35
LR-test	8789.12***	9000.81**	11145.12**	14789.14***
Tests				
SDM Vs SAR χ2-test	51.131***			
SDM Vs SEM χ2-test	120.98***			
SDM Vs SAC AIC	10250.4	12411.6	15434.85	16021.0
SDM Vs SAC BIC	10260.9	12417.6	15495.26	76765.1
Standard errors in parentheses				
p<0.05, ** p<0.01, *** p<0.001				

Note: +The control variables include the lags of the Logarithm of the Informal density one period t-1. The same for the native population density; the Log distance to the CBD identified as Unicentro in Bogota and two dummies; one that takes the value after 2018 when the highest peak of migration and the last one, a dummy that denotes Covid-19 after 2020. Errors are clustered by cadastrial sector. The Lr-Test of column 2 is perfomed contrastin SAM model against nesting models SAR-SEM-SAC.

The dependent variable; the density of formal an informal firm refers to number of firms by square kilometers. The independent variables are: population density which refers exclusively to native population by square kilometers; migration density is defined as Venezuelan migrants by birth place by square kilometers; the distance from the CBD is the Euclidian distance in square kilometers. Finally, it is included two dummy variables; one to reflect the highest migration peak in 2018, and the second to consider the Covid-19 pandemic in 2020. All the independent variables are lagged one period to avoid double causality.

The estimation results of the four models are shown in table 4. The local elasticities of the variable of interest, the logarithm of migrants per square kilometer: “Log Migration density (t-1)” are positive across all models, ranging from 0.115% in the SDM model, to 0.0712% in the SAC model.

The LR-test allows us to compare whether the SDM (which acts as the unrestricted model) is a better model than the SAR or SAC. All the LR-tests are significant, but the lowest is for the SDM as the best model. One can easily test the first null hypothesis whether the lag weighted coefficients are equal to cero and $\rho \neq 0$, thus the model is a SAR. Whereas if $\theta = -\beta\rho$, the model is a SEM. The Chi-2 tests rejects the first null hypothesis (that of comparing between the SDM and SAR models). Similarly, the Chi-2 tests for the second hypothesis (that of comparing between the SDM model against the SEM) goes for the SDM. Finally, the AIC-BIC Akaike Criteira rule out the SAC model.

The spatial autoregressive dependent variable collected by the Rho parameter is less than 1, revealing that the SDM model is the most suitable model and spatially non-explosive stable. Furthermore, evaluating the spatial autocorrelation error coefficient λ appears close to zero for the SEM and SAC models (columns 4 and 5), which again proves that there is not spatial autocorrelation in errors and reinforces the SDM model.

Regarding the local effects, a 1% increase in the previous migratory density is associated with a rise in the density of formal firms by 0,115% for the best model (SDM). For the global effects, 1% increase in immigration density for unit i is associated with a slightly decrease of the formal firms location in the closest neigborhoods by -0.124%. Basically, this occurs because the ability of firms to proliferate at spatial level is limited as the distance from any neighborhood increases, since this type of sector tends to agglomerate in specific areas of the city, in such a way that immigrants cannot influence increments at greater ranges of distance.

Table 4 also shows the average direct and indirect effects of migration on the proliferation of formal firms. Direct effects tell us, 1% increase in immigration in unit i, is associated with proliferation of formal firms in all neighborhoods in the city in about 0.809%. This magnitude also contains the feedback effect of this change from the neigborhoods coming back to original units i.

Regarding the indirect effects, 1% increase of immigrattion in neighboring areas is associated with null effects on neigborhood i (Table 4). This result is understandable, to the extent that, throughout the territory, the location of formal activities vanishes, while the location of migrants tends to extend towards the periphery. Couple with the fact that formality requires higher amounts of capital for migrants, limiting their effect on formality businesses.

The extended table of all control variables are shown in Appendix Table 3A, the first part of the table shows the non-spatial coefficients, in the second part, denoted by Wx, indicates that the variables are weighted by spatial weighted matrix.

In sum, we found evidence for a conservative extensive margin in formal firms associatted with migration, though, the spillover effects reveal, that in average, inmmigrants influence slighly, the pattern of proliferation of formal firms across the city. This is in part, because, the capacity of inmigrants to run businesses into the formal sector is limited due to legal restrictions and high entrepeneurship costs.

Now, the case of the model where the dependent variable corresponds to the density of informal firms is shown in Table 5. Once again, the LR-test is also significant in favor of the unrestricted SDM model. As observed in Table 5, a local increase of 1% in the density of migrants in neighborhood j is associated with a rise in the density of informal firms in j, by 0.911% for the SDM

Table 5. Estimation of the effects of Venezuelan migration on the informal business creation

	Depvar: Log Informal density			
Indepvar :Log Migration density (t-1)	(SDM)	(SAR)	(SEM)	(SAC)
Local effects	0.911***	0.902***	0.786***	0.123***
	(0.0314)	(0.0378)	(0.0632)	(0.032)
Wx: Global effects	0.679***			
	(0.00175)			
Direct effects	1.087***	0.8536		2.0189**
	(0.064257)	(0.0081238)		(0.822)
Indirect effects	6.894***	4.2540***		2.230***
	(0.296533)	(0.0399718)		(0.1054593)
total effects	7.982**	5.107**		4.248***
	(0.2964025)	(0.0480388)		(0.105331)
Spatial rho	0.287***	0.299***		0.281***
	(0.000167)	(0.000113)		(0.00356)
Lambda			0.099***	0.00105***
			(0.0000741)	(0.0000324)
Variance sigma2_e	8.356***	10.72***	12.72***	16.51***
	(0.014)	(0.301)	(0.102)	(0.014)
Time effects	Yes	Yes	Yes	Yes
All controls+	Yes	Yes	Yes	Yes
N	11,710	11,710	11,710	11,710
R-sq	0.22	0.14	0.21	0.11
adj. R-sq	0.17	0.10	0.19	0.10
LM-Test	4978.1***	5822.2***	6934.4***	7852.5***
Tests				
SDM Vs SAR χ2-test	101.12***			
SDM Vs SEM χ-test	90.89***			
SDM Vs SAC AIC	9250.4	10461.2	15420.15	15728.05
SDM Vs SAC BIC	9260.9	10471.4	15594.23	15732.02
Standard errors in parentheses				
* p < 0.05, ** p<0.01, *** p<0.001				

Note: +The control variables include the lags of the Logarithm of the Informal density one period t-1. The same for the native population density; the Log distance to the CBD identified as Unicentro in Bogota and two dummies; one that takes the value after 2018 when the highest peak of migration and the last one, a dummy that denotes Covid-19 after 2020. Errors are clustered by cadastrial sector. The Lr-Test of column 2 is perfomed contrastin SAM model against nesting models SAR-SEM-SAC.

model. The extended table of all control variables are shown in Appendix Table 4A.

Regarding the global effects of the migration variable, the higher the migration density in neighboring areas, a slightly positive competition effect is generated for informal companies in spatial unit j, since the elasticity is 0.679%. In other words, increments in immigration generate local and global positive effects on the creation of informal companies in all the city.

The autoregressive effects of the dependent variable collected by the Spatial Rho parameter, it is less than 1, showing that the SDM model is a spatially non-explosive stable model. The direct effects of migration found that on average a higher density of migrants in the city, generates an increase of 1.0876% in the density of informal firms in the city.

For the indirect effects, an increase in the immigrants concentration in all surrounding neighborhoods has a positive strong net effect on informal density of about 6.8949% in the city. This result is understandable, to the extent that, throughout the territory, the location of informal activities becomes intensive as the distance from the CBD rises, at the same time that, location of immigrants tends to locate towards the periphery. Thus, the two phenomena act jointly to the extent that, both are moving from the center to the periphery.

Also, the models show that there is no autocorrelation in the errors. In fact, the complementary SEM and SAC models that capture this autocorrelation, even though it is positive, is almost null for both estimates. Finally, all the tests for goodness of fit confirm again that SDM is the most suitable model.

In sum, the informal extensive margin of migration is slightly positive at local levels but it is stronger the contagious spillover effects across the city. Thus, where immmigrants arrive it ends up expanding the agglomeration of informal firms across the whole territory with a greater power than in the case of the formal entrepeneurial tissue.

6.1 Robustness checks

The previous estimations control for endogeneity by lagging all the independent variables one period behind. Here we will implement an instrumental variable strategy to reinforce the potential bias correction derived from endogeneity. We introduce a Bartik-shift-shared instrument proposed by Altonji and Card (1991) and Card (2001), based on the past settlement of immigrants.

The new estimations for the formal and informal regressions are shown in Appendix Tables 5A and 6A, all the tests continue to choose the SDM as the best model. These two tables show two kinds of regressions depending on using two types of instrumental variables. Regression labelled by “IV1” runs the first step estimation under a non-spatial fixed effects specification and regression labelled “IV2” runs the first step regression with a fixed effects spatial regression type SDM model. All the F-test of the first stage are higher than 50. It also was implemented the F Montiel-Plueger robust which is also above 50.

The Hausman test compares the initial lagged models presented in Table 4 with those of instrumental variables, where the null hypothesis claims that differences between them are not systematic. For the formal firms, appendix 5A shows that for all models (SDMIV1, SAC and SAR) the Hausman test does not reject the null hypothesis, except for the SDMIV2, (column 3 appendix Table 5A page 40). Meanwhile, comparing regressions with the two types of instruments (the non-spatial instrument and the spatial instrument type SDM structure) the lowest AIC-BIC favours the spatial one.

In the latter case (the spatial instrument), the local effects for formal firms are slightly bigger (0.224) compared with the original model (0.115), hence we were underestimating them. The global effects are twice small, so we would be overestimating them in the original model. For the direct effects, they remain positive and in magnitudes close to the first estimate (those in Table 4), but the indirect effects are smaller. Hence, instrumenting with a spatial IV slightly adjusts down the effects of migration on the expansion of formal firms but the direction becomes unchanged.

In the case of informal firms (Appendix Table 6A) the Hausman test is significant only for SDMIV2 and SAR IV2 models. However, comparing the lowest AIC-BIC criteria, the chosen model is the SDMIV2. Local effects area slightly small (0.863) compared the initial model (0.911). Global effects also diminished (0.486 against 0.679 of the initial model). Regarding the direct effects remained around same value (1.087 to 1.16), but indirect effects slightly go down (6.89 to 5).

In sum, the instrumental variable regression keeps the better goodness of fit and adjust downward the estimates. Also, it reinforces the fact that the association of formal firms and inmmigration is conservative not only at the local level of the neigborhood but also by spreading across the territory. For the informal firms, it keeps the higher spillover effects.

We continue to examine whether the estimates change when splitting data by firm size and economic sector. To perform this exercise, we have selected the SDMIV2 structure. The results for formal firms are shown in Appendix 7A Part1, where, the most appropriate model is still the SDM. Here, the effects of migration are null in large and medium-sized formal firms, as expected, as this type of economic units require a significant level of capital and legal regulations to operate.

Meanwhile the effects remain slightly positive for microenterprises and small formal firms where the requested capital investment is less demanding for entrepreneurs. On the other hand, the direct and indirect effects of migration are significant and 8 times higher in the case of informal microenterprises compared to formal ones (Appendix 7A Part2). The effect is also significant for the case of small informal firms, as it was expected.

Regarding to the economic sector, we aim to examine in which activities firms expansion predominates most. Here we select a first group of activities that are characterized to be mobile, that is, are able to relocate or expand when facing a labour shock, especially low-skill immigrants. More mobile activities such as trade, manufacturing as well as low value-added services or low skill intensive services (arts, recreation; accommodations and food services) and building. There are others less mobile, such as high-value services (information; finance; professional services; and management) and high-skill intensive that are unlikely to respond to labour supply shocks.

As observed in Appendix 7A Part 3, the local, direct and indirect effects are higher in the trade sector and low value-added formal services. As expected, these are activities with greatest capacity to absorb low skilled labour shocks. Instead, in formal high-services, the local and direct effects are almost nil. In the case of the informal sector (Appendix 7A Part 4), results are similar; informality reacts better to migratory shocks in the most mobile economic sectors.

For the high skilled services (financial and real estate administrative support services), the best model is SEM. Partly because in these sectors the level of informality is very low and because of the high concentration of these services. in very specific areas like the CBD, smaller spillover effects are present. These are exclusive services to the best-qualified urban areas.

7 Spatial Heterogeneity Analysis

In the previous sections, we estimated the global association of migration with the proliferation of firms. We also found, timid coefficients at local and global effects. It opens the concern whether in some parts of the city the positive and negative association gets offset. That is why, it is needed to examine the spatial heterogeneity by leading the same previous specifications into a Geographic Weighted Regression setting (with the same spatial instrumental variable at first stage).

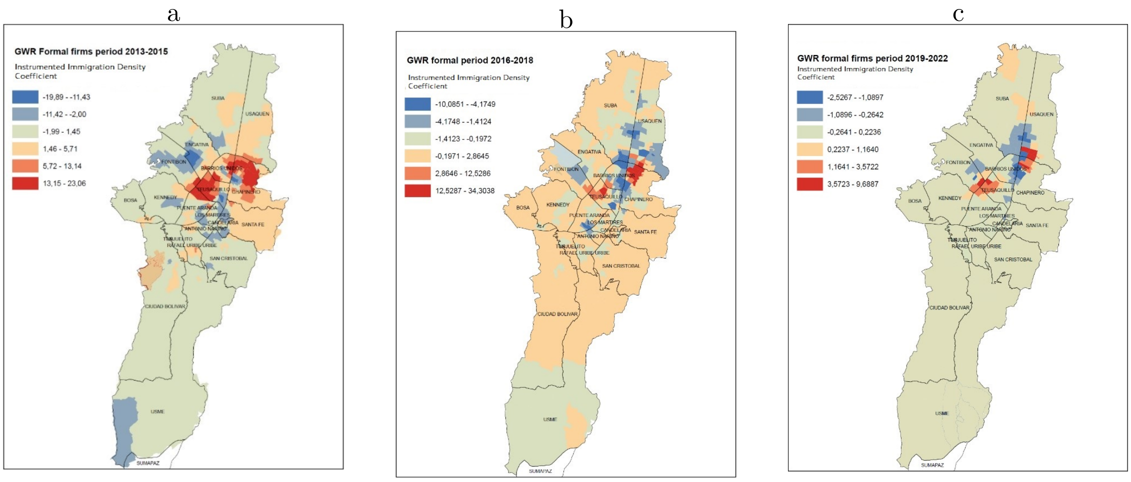
As it was observed in the empirical description of firms location, the spatial fragmentation is evident. Formal firms are concentrated in the CBD, while informal firms in the south west of the city. This means that the spatial behaviour of the economic activities presents a heterogeneous pattern; the more formal ones tend to locate on the CBD, while more informal are located somehow randomly in the space. When spatial heterogeneity is present, the parameters functional form may vary by location. To capture this heterogeneity, Fotheringham et al (1998) and Brudson et al (1998) proposed the method of Weighted Regressions in order to obtain an estimate of the vector of parameters for different observations by location. GWR is implemented in ArcGIS.

The first step is to apply a test of stationarity for confirming whether the association migration-firms varies along the space that justifies the use of spatial regression methods. In appendix Figure

2A, the significance tests for non-stationarity of the estimated coefficients show that the relationship between all the covariables changes significantly over the space (p-value<0.05). The test of the bandwidth suggests that the Geographically Weighted Regression model is a significantly better model for this data than the global linear regression model

With GWR analysis we pretend to examine whether the relationship between intra-period immigration and formal/informal firms is consistent across the study area. Which are the districts where migration is stronger associated with creation or destruction of formal/informal firms.

Map 4. GWR regression coefficients maps for formal firms model

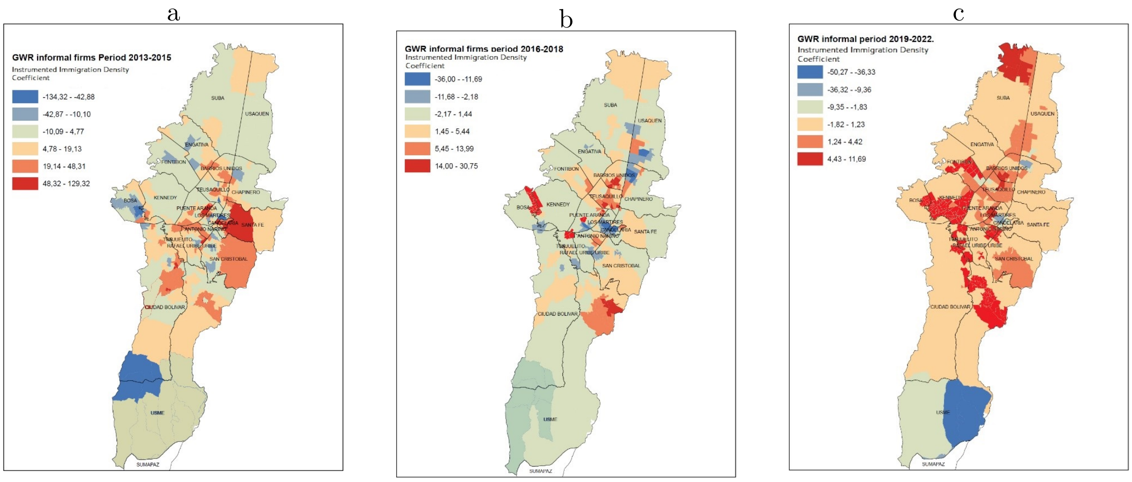


To perform this analysis, we aggregate all the observations for each variable in three time subperiods; The first one from 2013-2015, the second from 2016-2018 and the third from 2019-2022. In order to support this temporal partition, we verify whether the GWR coefficients remain stable over time. So, we follow the methodology of Kopczewska et al (2021), in which, the GWR regression was performed for each year individually, then, we cluster the GWR coefficients per year by applying the k-means clustering technique. Appendix Figure3A shows that the optimal level of clusters for all years is 9. Subsequently, we calculate the Rand index that tells us of the similarity or differences between comparisons of pairs of years followed. If the Rand index is close to 1 then the clustered coefficients are stable over time. If it is close to zero, then they vary significantly between pairs of years.

The Rand index is plotted in tables (see Appendix Figure 4A). The elements above the diagonal show the differences between pairs of years. The numbers from 1 to 9 denote 9 years and the year 2022 is left as the base year. In the case of formal firms, what is observed in figure 1A is that in fourth year, corresponding to the pair of years 2016-217 exhibits a drop in the Rand index. Here there is a slight break in the stability of the coefficients. So, the first subperiod should be partitioned in that moment. Likewise, in the 2018-2019 pair case. This confirms that the separation into three annual groups of map 4 works. For the case of informal firms, the same structure is maintained (see Appendix Figure 5A).

The set of Maps 4 shows the GWR coefficients for formal firms. Map 4a refers to the first period 2013-2015, Map 4b refers to the second period 2016-2018 and Map 4c to period 2019-2022. It shows how the regression coefficient change in time across neighborhoods. In these maps, only regression

Map 5. GWR regression coefficients maps for informal firms model



Note: p-value<0.05

coefficients significant at 5% significance (p-value<0.05) have been chosen. The coefficients that have a positive magnitude are denoted in red, and in blue those with a negative sign.

The red polygons denote where the GWR model shows that Venezuelan migration had a greater association with formal firms proliferation. In the first period, these neighborhoods are the ones located in the central parts right near the CBD; For instance, Chapinero, Teusaquillo and Barrios Unidos. Instead, in the blue zones, to the west in the Fontibon area, Venezuelan Lagged migration is associated with significant reductions in formal firms. Venezuelan migration influenced much more the creation of formal firms than in any other period.

While, in the second and third periods, the capacity of Venezuelan migration to generate the greatest densification of formal units was limited to a smaller polygon of neighborhoods in the central areas. The GWR association of Venezuelan migration on the proliferation of informal firms is evidenced in the set of maps in Map 5.

A common trend is observed throughout the three subperiods: Venezuelan migration is positively associated with the proliferation of informal businesses outside the CBD towards the south east. The magnitude of the spatial coefficients is higher than in the formal case seen previously. Likewise, the association of Venezuelan migration with a greater concentration of informal firms increased over time, contrary to what happens with formal firms.

8 Conclusions

Literature about migration has focused on the effects of migration over wages and unemployment concluding that their impacts are null. If this is so, we have proposed that firms also react to labor supply shocks through the extensive margin as a compensatory mechanism. We are supported on several authors that the informal sector is so large and low-productive that it is more difficult for firms to react at intensive margin. Also, the extensive margin has been little analyzed in the urban context of a developing country.

We found evidence that the extensive margin mechanism holds in the case of Venezuelan migration in Bogotá-Colombia during the period 2013-2022. Firms in Bogotá respond to immigration by increasing the number of establishments within a city in order to accommodate the abundant supply of workers. Being a developing country, the expansion of firms is greater in the informal sector than in the formal one. In fact, applying the best specification of the Durbin Model-SDM, the local effects reveal that an increase in 10% of the density of Venezuelan migrants in a neighbourhood, generates an increase of 1.15% in the density of formal firms in the city.

This result gives rise to the design of public policies to understand the differential role of the formal and informal sector in generating income for the migrant and native population. If migration can have a positive association with the dynamism of formal businesses, the magnitude of this association is weak compared to the informal one. This is partly due to the legal restrictions that migrants face to start formal businesses and represent additional costs in firmsâ creation (migration documentation, registration costs, social security payments). For this reason, public investment programs should encourage the creation of formal firms through seed capital, business strengthening and formal access to credit.

Promoting incentives such as cost reduction of registering a company or tax discounts when a specific quota for migrant employment is satisfied or a subsidy for formal employment is met. Additionally, productive policies must address the phenomenon of informality as an integral element of economic activity in developing countries. For instance, the promotion for business creation even within the informal sector, mainly at the initial economic cycles. At initial stages of business (the first five years) the high fixed costs lead many companies to operate informally as a mechanism to achieve financial growth.

At this stage is needed to strengthening the access to credit and seed capital for informal companies aiming at consolidating them to become formal in the medium term. The second finding is the identification of positive spillover effects of Venezuelan migration over time and space in Bogota. The arrival of migrants does affect the creation of firms throughout the territory in a chained way. But these effects are stronger for the informal sector.

Regarding the spillover effects, the direct effects are positive, but timid in the creation of both formal firms (0.809%) and in informal firms (1.0876%) in the city. However, the indirect effects are null for the formal sector, while in the informal sector they are five times higher. In other words, migrants do contribute to encourage the expansion of informal units at the aggregate level for the entire city, but they do not manage to generate an aggregate effect on business formality. In this way, the urban structure counts on the ability of migrants to influence the dynamics of the location of firms.

This result is key in the design of urban territorial planning policies and tells the importance of strengthening employment centers (CBDS) with the provision of strategic infrastructure in those places where there is a greater migratory concentration, which potentially is an indirect impulse for business creation.

Thus, immigration does not have the same effects on business formality and informality throughout the city. Venezuelan migration has had influence to create formal firms in the central parts right near the CBD; Chapinero, Teusaquillo and Barrios Unidos. While, in the west Fontibon area, Venezuelan immigration is associated with informal and the outside part of the CBD, more towards the south east. In these areas, public resources for productive infrastructure could be addressed to encourage creation that exploits the presence of Venezuelan labour.

The third result is that immigration is associated to an increase in the number of relatively small establishments but has little relationship with the number of larger establishments. On the other hand, the direct and indirect effects of migration are significant and 8 times higher in the case of informal microenterprises compared to formal units. The effect is also significant for the case of small informal firms, as was to be expected.

Finally, the increase in establishments predominantly occurs in industries that are relatively mobile. Trade and low-value-added services have a greater capacity to absorb a labour shock, although the capacity is greater in the informal sector. But in high-services, the local and direct effects are almost nil as expected, although the global effects are relatively high. Therefore, entrepreneurship policies should focus on the informal sector of micro and small firms and on relatively mobile sectors in concrete locations in the city.

References

[1] Anderson, W. (2001). Oral history and migrant wage labor: sources of narrative distortion. Oral History Review, 28, 2, pp 1–20.

[2] Albrecht, J., Navarro, L., and Vroman, S. (2009). The Effects of Labour Market Policies in an Economy with an Informal Sector. The Economic Journal 119, no. 539. 1105 –1129.

[3] Altonji, J., and Card, D. (1989). The Effects of Immigration on the Labor Market Outcomes of Natives. NBER Working Paper No. 3123.

[4] Altonji, J., and Card., D. (1991).The Effects of Immigration on the Labor Market Outcomes of Less–skilled Natives. Chapter in Immigration, trade and the Labor Market, 1991, pp. 201–234 from National Bureau of Economic Research, Inc.

[5] Amuedo–Dorantes, C. and S. Pozo. (2011). Accounting for Remittances and Migration on Children’s Schooling. World Development, 38(12), pp.1747 –59.

[6] Altindag, O., Bakis, O., and Rozo, S. V. (2020). Blessing or burden. Impacts of refugees on businesses and the informal economy. Journal of Development Economics, 146, 102–490.

[7] Antrop, M. (2005). Why landscapes of the past are important for the future?. Landscape and Urban Planning, 70(12): 21 –34.

[8] Accetturo, A., Francesco, M., Sauro, M., and Olivieria, E. (2014). Don’t stand so close to me: The urban impact of immigration. Regional Science and Urban Economics, Elsevier, 45, 45 –56.

[9] Audebert, C. (2013). Spatial Strategies of Haitian Businesses in the Diaspora: The Case of Metropolitan Miami (2001–2009). Journal of Haitian Studies, 19(1), 217 –234.

[10] Azoulay, P., Benjamin F., Kim, D., and Miranda., J. (2021). Immigration and Entrepreneurship in the United States. American Economic Review: Insights, forthcoming.Vol. 4(1), pages 71-88, March.

[11] Bahar, D., Iba nez, A., Rozo, S. (2020). Give Me Your Tired and You Poor: Impact of a Large–Scale Amnesty Program for Undocumented Refugees. Journal of Development Economics, Elsevier, vol. 151(C).

[12] Baghdadi, L. (2005). Mexico–U.S. Migration: Do Spatial Networks Matter?. Université Paris I., mimeograph, 2005.

[13] Bell, E., and Nkomo, S. (2001) Our Separate Ways: Black and White Women and the Struggle for Professional Identity. Berkeley, CA: Third Woman Press.

[14] Bhagwati, J., and Srinivasan, T.(1983). On the choice between capital and labour mobility, Journal of International Economics, Elsevier, vol. 14(3 –4), pages 209 –221, May.

[15] Borjas, G. (2003). The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. The Quarterly Journal of Economics, 118(4), 1335 –1374.

[16] Borjas, G., and Katz, L. (2007). The evolution of the Mexican born workforce in the United States.Pages 13-56, National Bureau of Economic Research, Inc.

[17] Brunsdon C., Fotheringham A., and Chambers M. (1998). Measuring spatial variations in relationships with geographically weighted regression.Advances in Spatial Science. Springer, Berlin, Heidelberg. $https://doi.org/10.1007/978-3-662-03499-64$

[18] Cavailhés, J., Peeters, D., Sekeris, E., and Thisse, J.F. (2004). The periurban city: why to live between the suburbs and the countryside?, Regional Science and Urban Economics, 34, pp. 681–703

[19] Crasto, T., Reguant, M. (2017). Perceptions of Venezuelan migration: Causes, Spain as a destination, Return expectations. Migraciones, 133 –163.

[20] Card, D.(2001). Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration.Journal of Labor Economics, 19(1), pp. 22 –64.

[21] Card, D. (2007). How Immigration Affects U.S. Cities. Center for Research and Analysis of Migration Discussion Paper Series No 11/07.

[22] Cagetti, M., and Mariacristina De Nardi (2006). Entrepreneurship, Frictions, and Wealth. Journal of Political Economy 114.5, pp. 835 –870.

[23] Card, D., and Lewis, E., (2007). The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts. In Mexican Immigration to the United States, edited by George Borjas, National Bureau of Economic Research Conference Report, Cambridge, MA.

[24] Combanouse, F., Degilhem, T. (2019). Une approche novatrice du marché du travail urbain de Bogota:le prisme de la qualité de l‘employ. Pages 283–316. Revue d’Economie Régionale et Urbaine.

[25] Constant, A (2014). Do migrants take the jobs of native workers?. IZA World of Labor. 10 doi: 10.15185/izawol.10.

[26] Constant, A., Zimmermann, K. (2006). The Making of Entrepreneurs in Germany: Are Native Men and Immigrants Alike?. Small Bus Econ 26, 279–300.$https://doi.org/10.1007/s11187-005-3004-6$

[27] Dustmann, C., and Glitz, A. (2011). Migration and education. Handbook of the Economics of Education. Vol. 4. Elsevier. 327 –439.

[28] Delgado–Prieto, L. (2022). Immigration, wages, and employment under informal labor markets. Working papers 96, Red Investigadores de EconomÃa.

[29] Eckstein, S., and Peri, G. (2018). Immigrant Niches and Immigrant Networks in the U.S. Labor Market. Journal of the Social Sciences, 4(1), 1 –17.

[30] Edo, A., and Rapoport, H, (2019). Minimum wages and the labor market effects of immigration, Labour Economics, Elsevier, vol. 61(C).

[31] Ellis, M., Wright, R., and Parks, V. (2004). Work together, live apart?. Geographics of racial and ethnics segregation at home and work. Annals of the Association of American Geographers, 94(3), 620 –637.

[32] Ellis, M., Wright, R., and Parks, V. (2007). Geography and the inmigrant division of labour. Economic Geographic, 83(3), 221 –230.

[33] Elhorst, J. P. (2010). Applied spatial econometrics: raising the bar. Spatial economic analysis 5(1): 9 –28.

[34] Evans, D., and Jovanovic, B. (1989). An estimated model of entrepreneurial choice under liquidity constraints. Journal of Political Economy, 97, 808 –827. doi:10.1086/261629.

[35] Fairlie, R., Meyer, B. (2003). The Effect of Immigration on Native Self–Employment, Journal of Labor Economics, 21, (3), 619 –650.

[36] Fischer M., Getis A. (eds.) Recent developments in Spatial Analysis, Springer–Verlag, London.

[37] Flores– Fillol, R., Garcia– López, M., and Nicolini, R. (2016). Organization of land surrounding airports: The case of the aerotropolis. Land Economics, 92(1), 57 –81.

[38] Fotheringham, A., Stewart, W., and Wei, K. (2017). Multiscale geographically weighted regression (MGWR). Annals of the American Association of Geographers 107.6: 1247 –1265.

[39] Freitez, A. (2011). La emigración desde Venezuela en la última década. Temas de Coyuntura, 63, 11 –38.

[40] Friedberg, R., Hunt, J. (1995). The Impact of Immigrants on Host Country Wages, Employment and Growth. Journal of Economic Perspectives, 9:2, 23 –44.

[41] Friedberg, M. (2001). The Impact of Mass Migration on the Israeli Labor Market. The Quarterly Journal of Economics, Vol. 116, No. 4, pp. 1373 –1408.

[42] Garcia L,. Nicolini, R., and Roig J. (2021). Urban spatial structure in Barcelona (1902–2011): Immigration, spatial segregation and new centrality governance (Open Access) Applied Spatial Analysis and Policy, vol. 14(3), pp. 591–629.

[43] Gianmarco, I., Ottaviano, G. (2012). Rethinking the Effect of Immigration on Wages. Journal of the European Economic Association, 10(1), 152–197.

[44] Glitz, A. (2012). The Labor Market Impact of Immigration: A Quasi–Experiment Exploiting Immigrant Location Rules in Germany. Journal of Labor Economics, 30(1), pp.175–213.

[45] Gonzalez, L., Ortega, F. (2008). How Do Very Open Economies Adjust to Large Immigration Flows? Recent Evidence from Spanish Regions. IZA Working Paper 3311.

[46] Grogger, J., Hanson, J. (2008.). Income Maximization and the Selection and Sorting of International Migrants. NBER Working Paper No 13821.

[47] Gross, D., and Schmitt, N. (2003). The role of cultural clustering in attracting new immigrants. Journal of Regional Science, 43(2), 295–318.

[48] Hall, S. (2004). Divided city: the crisis of London. openDemocracy. Retrieved from $https : //www.opendemocracy.net/arts - multiculturalism/article_2 191.jsp[GoogleScholar]$.

[49] Hatziprokopiou, H., Frangopoulos, Y., and Montagna, N. (2016). Migration and the city Diversity, Migrant Economies and Urban Space. City, 20(1), 52–60.

[50] Kaplan, D., and Li. (2006). Landscapes of the Ethnic Economy. United Kingdon. Rowman and Littlefield.

[51] Kerr, S., Kerr, W. (2011). Economic Impacts of Immigration: A Survey. Finnish Economic Papers, Finnish Economic Association, 24(1), 1–32.

[52] Koike, S (2015). Urban structure, labor market, informal employment and gender in Mexico City. Dissertation thesis. Universitat Autonoma de Barcelona.

[53] Kopczewska, K., Cwiakowski, Piotr.(2021). Spatio-temporal stability of housing submarkets. Tracking spatial location of clusters of geographically weighted regression estimates of price determinants. Land use Policy. Volume 103, 105292.

[54] Kleemans, M., and Magruder, J. (2018). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. The Economic Journal, 128(613):2032–206.

[55] Lacruz, T. (2006). Balance sociopolÃtico: una ciudadanÃa social inacabada. Balance y perspectiva de la polÃtica social en Venezuela, 112– 184. Available in https://www.redalyc.org/pdf/403/40316176004.pdf.

[56] Leamer, E., and Levinsohn, J. (1995). International trade theory: the evidence. Handbook of international economics, 3, 1339–1394.

[57] LeSage, J., Kelley, P. (2010). Spatial econometric models. In: Fischer MM, Getis A (eds), Handbook of applied spatial analysis: Software tools, methods and applications. Springer, Berlin.

[58] LeSage, J., Pace, R.(2009). Introduction to spatial econometrics. Chapman and Hall/CRC, Boca Raton FL.

[59] LeSage, J., and Pace, R. (2001). Spatial dependence in data mining. In Data mining for scientific and engineering applications (pp. 439–460). Springer, Boston, MA.

[60] Lewis, E. (2003). Local open economics within the U.S. how do industries respond to immigration?, Federal Reserve Bank of Philadelphia Working Paper.

[61] Lewis, E. (2005). Immigration, Skill Mix, and the Choice of Technique, Federal Reserve Bank of Philadelphia Working Paper 05–08.

[62] Lewis, E., (2011). Immigration, Skill Mix, and Capital Skill Complementarity.Quarterly Journal of Economics, 126(2), pp. 1029–1069.

[63] Lewis, E., (2013). Immigration and Production Technology.Annual Review of Economics 5, pp. 165–191.

[64] Li, W. (1998). Los Angeles Chinese etnoburb; from ethnics service sector to global economy outpost. Urban Geography, 19(6), 502–517.

[65] Longhi, S., Nijkamp, P., and Poot, J. (2005). A Meta–Analytic Assessment of the Effect of Immigration on Wages, Journal of Economic Surveys, Vol. 19, No. 3, pp. 451–477.

[66] Longhi, S., Nijkamp, P., and Poot, J. (2010). Joint impacts of immigration on wages and employment: review and meta–analysis. J Geogr Syst, 12, 355387.

[67] Lebow, J. (2021). Immigration and Occupational Downgrading in Colombia.Thesis dissertation University of Duke.

[68] Logan, J. R., Alba, R., and Zhang, W. (2002). Immigrant enclaves and ethnic communities in New York and Los Angeles. American Sociological Review, 67(2), 299–322.$https://doi.org/$10.2307/3088897.

[69] Lozano, H., Urbina, J., and Dominguez,A. (2021). A. Estructura urbana y cambio en el uso del suelo en Bogotá entre 2007 y 2017. urbe, Rev. Bras. Gest. Urbana [online], vol.13 [cited2022–04–05], e20200192. Available from: $http://old.scielo.br/scielo.php?script = sciarttextpid = S2175 - 33692021000100284lng = ennrm = iso$. Epub Nov 29, 2021. ISSN 2175–3369.

[70] Morales, Z., Bonilla,M., Hermida L., Florez,. D.(2020). The labor market of immigrants and non–immigrants evidence from the Venezuelan refugee crisis. Borradores de Economía; No. 1119.

[71] Muñiz, I., Sánchez, V. (2019). La decentralisation de la population et de l'emploi entraine–t–elle une réduction des distances domicile–travail ?. Le cas de la zone métropolitaine de la vallée de Mexico. Revue deconomie Regionale et Urbaine. 2–2019. 259–281. 10.3917–reru.192.0259.

[72] Nicolini, R., Roig, J. (2019). Regional and Urban economics in Latin America: Informality and economic development. Pages 197–2011. Revue dEconomie Régionale et Urbaine.

[73] Olney, W. (2013). Immigration and firm expansion. Journal of Regional Science, 53, 142–157.

[74] Otaviano, I,. and Peri, G. (2006). Rethinking the Effects of Immigration on Wages. NBER Working Papers.

[75] Ottaviano, I., Peri, G., Wright, G. (2013). Immigration, Offshoring and American Jobs. American Economic Review, 103 (5): 1925–59.

[76] Parks, V. (2004). Immigrant enclaves and ethnic niches: The gendered connection between residential and labor–market segregation in Los Angeles, Urban Geography, Vol. 25, 589–630.

[77] Peri, G., and Sparber, C. (2009), Task Specialization, Immigration, and Wages. American Economic Review, 1(3), pp. 135–169.

[78] Pischke J., and Velling J. (1997). Employment Effects of Immigration to Germany: An Analysis Based on Local Labor Markets. The Review of Economics and Statistics, 79(4), pp. 594–604.

[79] Posada, H., Moreno,A. (2019) Regional and Urban economics in Latin America: Informality and economic development. Pages 213–238. Revue dEconomie Régionale et Urbaine.

[80] Portes, A. (2000). Immigration and the Metropolis:Reflections on Urban History. Journal of International Migration and Integration, 1(2), 153–175.

[81] Sarker, A., and Unel, B. (2017). The Impact of Bank Expansion on Self–Employed Business Owners: Evidence from US States. No. 2017–06.

[82] Saiz, A. (2003). Room in the kitchen for the melting pot: Immigration and rental prices. Review of Economics and Statistics, 85(3), 502–521.

[83] Saiz, A. (2007). Immigration and housing rents in American cities. Journal of Urban Economics, 61(2), 345–371.

[84] Santamaria, J. (2020). When a Stranger Shall Sojourn with Thee: The Impact of the Venezuelan exodus in Colombia labor market. Thesis dissertation University of Minessota.

[85] Schoeni, R. (1997), The Effects of Immigration on the Employment and Wages of Native Workers: Evidence from the 1970s and 1980s. RAND Working Paper.

[86] Schwartz, A. (1973). Interpreting the Effect of Distance on Migration. Journal of Political Economy, 81(5), 1153–1169.

[87] Stoll, M., Raphael, S. (2000). Racial Differences in Spatial Job Search Patterns: Exploring the Causes and Consequences.Economic Geography 76(3): 201–223.

[88] Tribín– Uribe, A., Achyuta, A., Anzola, C., Ávila–Montealegre, Ó. I., Bonilla–Mejía, L., Castro–Fernández, J. C., Flórez, L. A., Grajales–Olarte, Á., Guarín–López, A., Hamann–Salcedo, F., Hermida–Giraldo, D., Khanna, G., Lasso–Valderrama, F. J., Medina–Durango, C., Melo–Becerra, L. A., M'éndez–Vizcaino, J. C., Morales–Zurita, L. F., Nyshadam, A., Ospina–Tejeiro, J. J., Otero–Cortés, A., Pulido, J. D., Ramos–Forero, J., Ramos–Veloza, M., Tamayo–Casta no, J. A., Velásquez, S. (2020). Migración desde Venezuela en Colombia: caracterización del fenómeno y análisis de los efectos macroeconómicos. Ensayos sobre Política Económica, (97), 1–74. $https://doi.org/$10.32468/$espe$.97.

[89] Topel, R. (1994). Regional Labor Markets and the Determinants of Wage Inequality. American Economic Association, 84(2), 17–22.

[90] Ulyssea, G. 2018. Firms, Informality and Development: Theory and Evidence from Brazil. American Economic Review 108, no. 8:2015–47.

[91] Ulyssea, G (2020) Informality: Causes and Consequences for Development. Annual Review of Economics, Vol. 12, pp. 525–546, Available at SSRN: https://ssrn.com/abstract=3669624 or http://dx.doi.org/10.1146/annurev–economics–082119–121914.

[92] Unel, B. (2018). Effects of Immigration on Native Entrepreneurship in the U.S, Departmental Working Papers, Department of Economics, Louisiana State University.

[93] Viseth, A. 2020. Immigration and Employment: Substitute versus Complementary Labor in Selected African Countries. IMF Working Paper 20/149, International Monetary Fund, Washington, DC.

[94] Wang, L. (2004). An investigation of Chinese immigrant consumer behaviour in Toronto, Canada. Journal of Retailing and Consumer Services 11 307–320.

[95] Waldinger, R., and Der–Martirosian, C.(2001). The Immigrant Niche: Pervasive, Persistent, Diverse in Strangers at the Gates: New Immigrants in Urban America, Roger Waldinger, ed. Berkeley, CA: University of California Press.

[96] Winter–Ebmer R., Zweimuller, J.(1996). Immigration and the Earnings of Young Native Workers. Oxford Economic Papers, 48, pp. 473–491.

[97] Wilson, F.D. (2003). Ethnic niching and metropolitan labour markets. Social Science Research, 32(3): 429–66.

[98] Wang, L. (2004). An investigation of Chinese immigrant consumer behaviour in Toronto, Canada. Journal of Retailing and Consumer Services 11 307 – 320.

[99] Wright, R., and Ellis, M. (2001). Immigrants, the Native–Born, and the Changing Division of Labor in New York City. In New immigrants in New York (pp. 81–110). Columbia University Press.

[100] Wilson, K., and Portes, A. (1980). Immigrant Enclaves: An Analysis of the Labor Market Experiences of Cubans in Miami. American Journal of Sociology, 86(2), 295–319.

[101] Wright, R., Ellis, M., and Parks, V. (2010). Inmigrants Niches and the intrametropolitan division of labour. Journal of Ethnics and Migration Studies, 36(7), 1033–1059.

APPENDIX

Table 2A.Capturing the urban structure for population of Bogota

	Depvar:Log Population residential density	
IndepVariables	(OLS)	(FE)
lndis_San Diego CBD	0.072***	0.1428***
	(3.45)	(5.21)
lndis Sagrado Corazon CBD	0.040	-0.030
	(1.32)	(1.05)
Lndis Unicentro_	0.1126**	0.169***
	(9.74)	(3.49)
lnDist Chapinero	-0.006	0.1151***
	(0.72)	(8.09)
lndist Chicó	0.055**	0.003
	(4.48)	(0.18)
lndis Modelia	-0.1331**	-0.1428
	(8.34)	(5.21)
lndis Salitre	-0.03	-0.085
	(1.56)	(3.63)
Constant	11.9***	11.43***
	(46.93)	(27.20)
R2 ADJ	0.101	0.321
N	10,785	10,785
Fixed effects district histtorical	No	Si
Year fixec effects	No	Si
F-statistics		63.42***
AIC	9.245***	8.236***
BIC	9.070***	7.320**

Robust errors at neighborhood level. ***significant 1%. **significant 5%. *Significant 10%. All independent variables are logarithms of distances in km2 from centroids of neighborhoods. The dependent variable is population by km2. For this estimation, we follow the proposal of Garcia et al (2019) to identify the main CBD through changes in the elasticity between the density of the population in Bogota and the physical distance from six selected points previously identified as candidates for CBDs. As Bogota exhibits a polycentric spatial distribution with possible centers: Unicentro, Chapinero, Chicó, Modelia, San Diego y Salitre. The equation to estimate is:

$$logD_{tj} = const + \alpha_0 lnx_{tj0} + \alpha_k \sum_{k=1}^{9} lnx_{tjk} + \mu_t + \delta_j + \epsilon_{tj}$$

In this equation,D_{tj} is the population (or community) density at time t in urban location j (district or neighborhood); x_{tj0} is the distance (in km) from the centroid of location j to a selected 0 point, e.g the traditional CBD; while x_{tjk} is the distance (in km) from the centroid of location j to a selected point k. The conclusion here, is that, the Unicentro subcenter presents a greater population centripetal force when fixed effects are introduced.

Figure 1A. Migration density cluster map

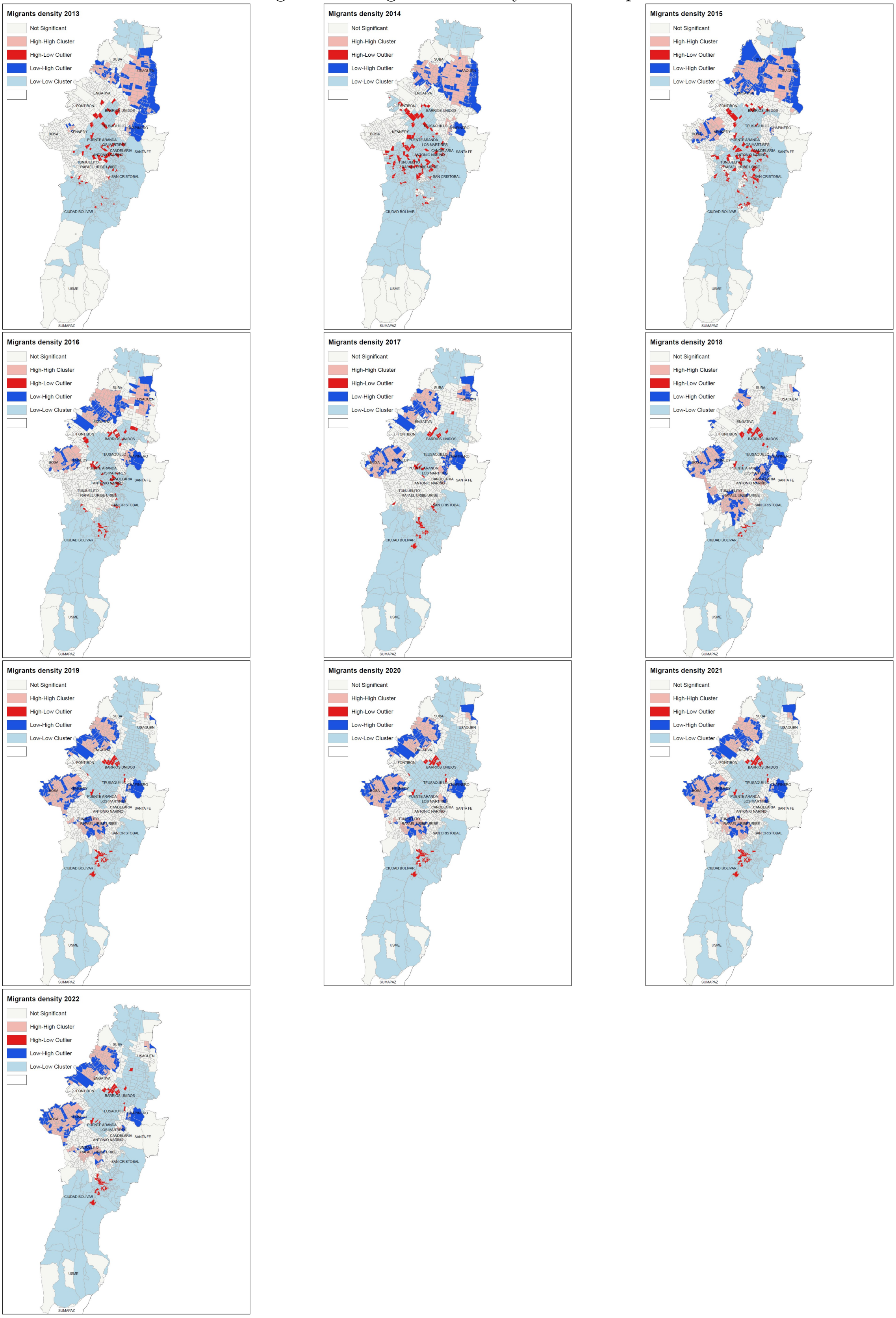


Table 3A. Extended estimation results for formal firms

	Depvar: Log Formal density			
IndVariables	(SDM)	(SAR)	(SEM)	(SAC)
Log Informal density (t-1)	0.380***	0.723***	-0.683*	0.467***
	(0.064)	(0.007)	(0.166)	(0.021)
Log population density (t-1)	-0.339***	-0.414***	-0.773***	0.0059
	(0.088)	(0.00533)	(0.207)	(0.029)
Log Migration density (t-1)	0.115***	0.136***	0.157***	0.0712***
	(0.00924)	(0.00891)	(0.00963)	(0.00832)
Log distance to CBD Unicentro	-0.766***	-0.773***	-0.707***	-0.824***
	(0.0160)	(0.0161)	(0.016)	(0.016)
Dummy_2018	0.324**	0.218	0.251*	-0.089*
	(0.177)	(0.150)	(0.103)	(0.038)
Dummy_pandemic 2020	-4.74**	-6.211**	-6.334**	-5.0279***
	(0.170)	(0.139)	(0.100)	(0.034)
Time effects	Yes	Yes	Yes	Yes
Wx				
Log Informal density (t-1)	0.374***			
	(0.0199)			
Log population density (t-1)	-0.138***			
	(0.0016)			
Log Migration density (t-1)	-0.124***			
	(0.0666)			
Log distance	-0.251***			
to CBD Unicentro	(0.0119)			
Dummy_2018	0.226			
	(0.381)			
Dummy_pandemic 2020	-2.361**			
	(0.00829)			
Time effects	Yes	Yes	Yes	Yes
Spatial rho	0.499***	0.499***		0.199***
	(0.000112)	(0.000111)		(0.000364)
lambda			0.0957***	0.00157***
			(0.0000358)	(0.000101)
Variance sigma2_error	4.966***	20.60***	36.13***	43.32***
	(0.0650)	(0.0406)	(0.866)	(0.567)
N	11,710	11,710	11,710	11,710
R-sq	0.26	0.263	0.28	0.43
adj. R-sq	0.22	0.23	0.18	0.35
LR-test	8789.12***	9000.81**	11145.12**	14789.14***
AIC	10250.4	12411.6	15434.85	16021.0
BIC	10260.9	12417.6	15495.26	16765.1
Standard errors in parentheses				
p<0.05, ** p<0.01, *** p<0.001				

Table 4A.Extended estimation results for informal firms

	Depvar: Log Informal density			
IndVariables	(SDM)	(SAR)	(SEM)	(SAC)
Log Formal density(t-1)	0.395***	0.391***	0.385***	0.339***
	(0.00765)	(0.00764)	(0.00759)	(0.00736)
Log population density(t-1)	0.262*	0.256*	0.231	0.0719
	(0.0124)	(0.0124)	(0.0122)	(0.0120)
Log Migration density (t-1)	0.911***	0.902***	0.786***	0.123***
	(0.0314)	(0.0378)	(0.0632)	(0.032)
Log distance to CBD Unicentro	0.145***	0.155***	0.144***	0.1272
	(0.019)	(0.011)	(0.012)	(0.012)
Dummy 2018	-0.0634	-0.109	-0.141	-0.0740*
	(0.152)	(0.132)	(0.123)	(0.0340)
Dummy pandemic 2020	-0.0563***	-0.156**	-0.223**	-0.114***
	(0.144)	(0.123)	(0.119)	(0.0294)
Time effects	Yes	Yes	Yes	Yes
Wx				
Log Formal density(t-1)	0.806***			
	(0.00376)			
Log population density(t-1)	0.162***			
	(0.0125)			
Log Migration density(t-1)	0.679***			
	(0.00175)			
Log distance to CBD Unicentro	0.110***			
	(0.0171)			
Dummy 2018	0.397			
	(0.322)			
Dummy pandemic 2020	0.107			
	(0.160)			
Time effects	Yes	Yes	Yes	Yes
Spatial				
rho	0.287***	0.299***		0.2808***
	(0.000167)	(0.000113)		(0.00356)
lambda			0.0993***	0.00105***
			(0.0000741)	(0.0000324)
Variance				
sigma2_e	8.356***	10.72***	12.72***	16.51***
	(0.014)	(0.301)	(0.102)	(0.014)
N	11,710	11,710	11,710	11,710
R-sq	0.22	0.14	0.21	0.11
adj. R-sq	0.17	0.10	0.19	0.10
LR-Test	4978.1***	5822.2***	6934.4***	7852.5***
SDM Vs SAC AIC	9250.4	10461.2	15420.15	15728.05
SDM Vs SAC BIC	9260.9	10471.4	15594.23	15732.02
Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001				

Table 5A. Instrumental Variables Estimation for immigration and formal firms association.

Variable	SDM IV1	SDM IV2	SAR IV1	SAR IV2	SEM IV1	SEM IV2	SAC IV1	SAC IV2
Main								
Log Informal density(t-1)	1.297***	3.413***	7.802***	6.753***	-1.370***	-0.208***	â0.233***	0.219***
	(0.174)	(0.094)	(0.367)	(0.134)	(0.012)	(0.003)	(0.007)	(0.007)
Log Migration density** (IV non spatial)	2.424***		3.621***		3.596***		0.558***	
	(0.436)		(0.858)		(0.027)		(0.017)	
Log Migration density** (IV spatial)		0.224***		0.250		0.118***		0.033
		(0.005)		(0.002)		(0.001)		(0.001)
Log population density(t-1)	-1.147***	-0.225	-6.806***	-3.986***	-0.331***	-0.309***	-0.079***	0.086***
	(0.139)	(0.128)	(0.193)	(0.097)	(0.005)	(0.002)	(0.0024)	(0.008)
Dummy 2018	0.147	0.401	0.787**	0.213***	1.883***	0.567***	0.345***	0.315*
	(0.182)	(0.219)	(0.107)	(0.967)	(0.022)	(0.016)	(0.0113)	(0.149)
Dummy pandemic 2020	-0.230	-1.016	-5.621***	-22.469***	-1.143***	-0.394***	-0.214***	-0.775
	(0.181)	(0.219)	(1.00)	(0.973)	(0.016)	(0.0170)	(0.008)	(0.149)
Wx								
Log Informal density(t-1)	0.321***	0.177***						
	(0.0021)	(0.0018)						
Log Migration density** (IV non spatial)	-0.403***							
	(0.004)							
Log Migration density** (IV spatial)		-0.066***						
		(0.0015)						
Log population density(t-1)	-0.167***	-0.079***						
	(0.001)	(0.00091)						
Dummy 2018	0.273***	0.062***						
	(0.012)	(0.0146)						
Dummy pandemic 2020	-0.512***	-0.126***						
	(0.0124)	(0.014)						
Spatial rho								
	0.298***	0.298***	0.298***	0.298***			0.0061***	0.0062***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)			(0.002)	(0.0004)
lambda					0.09565***	0.0956***	0.0956***	0.00626***
					(0.001)	(0.001)	(0.001)	(0.001)
Variance sigma2_e	17.049***	15.16***	86.91***	97.77***	24.764***	53.163***	5.018***	1.335***

continue...

Variable	SDM IV1	SDM IV2	SAR IV1	SAR IV2	SEM IV1	SEM IV2	SAC IV1	SAC IV2
Direct Effects								
Log Informal density(t-1)	0.138***	0.421***	0.593***	0.541***			0.237***	0.226***
	(0.448)	(0.119)	(0.234)	(0.203)			(0.0072)	(0.0078)
Log Migration density** (IV non spatial)	1.237**		2.740***				0.567***	
	(0.402)		(0.108)				(0.017)	
Log Migration density** (IV spatial)		0.824***		0.8417***				0.684
		(0.0054)		(0.00113)				(0.001)
Log population density(t-1)	-0.497**	-0.327***	-0.510***	-0.112***			-0.081***	-0.089***
	(0.211)	(0.172)	(0.201)	(0.420)			(0.0025)	(0.008)
Dummy 2018	1.016***	0.758***	0.214**	1.7202***			0.350***	0.326*
	(0.844)	(0.189)	(0.969)	(0.641)			(0.0113)	(0.151)
Dummy pandemic 2020	-1.888**	-0.613	-0.413***	-0.179***			-0.217***	-0.0089
	(0.884)	(0.109)	(0.168)	(0.678)			(0.00794)	(0.150)
Indirect effects								
Log Informal density(t-1)	-0.115***	-0.2501***	0.906***	0.763***			-3.419***	-0.7007
	(0.4093)	(0.109)	(0.212)	(0.185)			(0.171)	(0.406)
Log Migration density** (IV non spatial)	0.101**		0.423***				0.817***	
	(0.366)		(0.985)				(0.388)	
Log Migration density** (IV spatial)		0.398***		0.228***				0.121
		(0.511)		(0.106)				(0.347)
Log population density(t-1)	-0.056***	-0.032***	-0.799***	-0.158***			-1.165***	-2.783
	(0.1933)	(0.1553)	(1.183)	(0.382)			(0.067)	(0.665)
Dummy 2018	0.447***	0.282***	0.322	2.405***			0.506**	1.034
	(0.747)	(0.169)	(0.875)	(0.583)			(0.268)	(0.968)
Dummy pandemic 2020	-0.114***	-0.927	-0.670***	-0.253***			-3.135***	-0.1225
	(0.792)	(0.962)	(0.154)	(0.617)			(0.185)	(0.52)
Total Effects								
Log Informal density(t-1)	0.023***	0.171***	1.499***	1.304***			-3.182***	-0.474
	(0.0524)	(0.1099)	(0.227)	(0.185)			(0.1704)	(4.06)
Log Migration density** (IV non spatial)	1.338***		3.162***				1.383***	
	(0.0978)		(0.100)				(0.383)	
Log Migration density** (IV spatial)		1.222***		1.070***				0.806
		0.0545		0.0516				0.0356
Log population density(t-1)	-0.553***	-0.361***	-1.309***	-0.269***			-1.083***	-2.873
	(0.0303)	(0.0267)	(0.183)	(0.379)			(0.0676)	(1.668)
Dummy 2018	1.464***	1.040*	0.536*	4.125***			0.857***	1.361
	(0.319)	(0.413)	(0.221)	(0.609)			(0.268)	(0.976)
Dummy pandemic 2020	-2.00***	-1.541	-1.084***	-0.433***			-3.353***	-0.132
	0.315	0.383	0.229	0.635			0.186	0.534

continue...

Variable	SDM IV1	SDM IV2	SAR IV1	SAR IV2	SEM IV1	SEM IV2	SAC IV1	SAC IV2
Lr-test	9214***	8346.2***	9148.8***	9149.0***	9566.3***	9374.5***	9866.8	9801.9***
SDM Vs SAR χ2-test	20.35***	21_34***						
SDM Vs SEM χ2-test	102.54**	101.02**						
SDM Vs SAC AIC	12991.4	8934.3	17186.1	9042.1	19676.0	9921.3	19676.0	10922.3
SDM Vs SAC BIC	11923.1	7657.3	12334.8	7987.4	15304.3	8992.5	15304.3	10982.1
Hausman test: IV non Spatial migrants 2005 vs intial model	0.18		0.22		0.43		0.35	
F First Stage IV1	104.6		104.6		104.6		104.6	
Montiel Pflueger Robust IV1	88.23		88.23		88.23		88.23	
IV Spatial migrants 2005 vs intial model		260.40**		30.2		20.1		12.3
F First Stage IV2		197.61		197.61		197.61		197.61
Montiel Pflueger Robust IV2		96.23		96.23		96.23		96.23
Time effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.56	0.52	0.72	0.57	0.68	0.47	0.43	0.46
R2 adj	0.46	0.5	0.68	0.51	0.63	0.4	0.38	0.45
N	12.881	12.881	12.881	12.881	12.881	112.881	12.881	12.881

Note: “IV1” refers to the first stage non-instrumental regression. While “IV2” refers to the first stage spatial instrumental regression. +The control variables included the lags of the Logarithm of the formal density of establishments one period behind t-1. Also, for the logarithm of the native population density; the Log distance to the CBD (Unicentro) and two dummies (one that takes the value after 2018 when the highest peak of migration and the last one, a dummy that denotes Covid-19 after 2020). Errors are clustered by neighborhood. The Lr-Test of column 2 is performed contrasting SAM model against nesting models SAR-SEM-SAC. The Chi-2 test allows to compare the SDM model against a SAR or SEM model. While the AIC-BIC criteria allow to rule out the SAC model in favor of SDM model. The variable distance to the CBD is also included.

Table 6A.Instrumental Variables Estimation for immigration and informal firms association.

Variable	SDM IV1	SDM IV2	SAR IV1	SAR IV2	SEM IV1	SEM IV2	SAC IV1	SAC IV2
Log Formal density(t-1)	0.4933***	0.3547***	0.932***	0.4733***	0.2510***	0.1214***	0.1602***	0.569***
	(3.57)	(5.40)	(1.22)	(2.93)	(3.06)	(7.84)	(7.36)	(3.62)
Log Migration density** (IV non spatial)	1.4198***		1.6878***		1.113***		1.055***	
	(0.53)		(1.50)		(45.67)		(16.60)	
Log Migration density** (IV spatial)		0.8632***		0.8012***		0.7226***		0.139***
		(0.79)		(0.46)		(0.79)		(0.50)
Log population density(t-1)	1.3974***	1.7445***	2.374***	2.840***	0.2042***	0.5028***	0.6803***	6.883***
	(10.18)	(23.43)	(3.42)	(2.44)	(1.01)	(1.96)	(3.73)	(2.49)
Dummy_2018	7.8076***	6.0308***	7.974***	7.1879***	0.8399***	0.3280***	1.1541***	5.4119***
	(6.07)	(4.75)	(7.53)	(9.48)	(3.82)	(4.56)	(7.50)	(3.65)
Dummy_pandemic 2020	-8.68***	-7.38***	-7.98***	-9.35***	-0.30***	-0.60***	- 0.34***	-9.41***
	(6.65)	(5.81)	(2.00)	(3.31)	(4.64)	(8.56)	(5.81)	(4.19)
Wx								
Log Formal density(t-1)	0.7569***	0.756***						
	(2.62)	(6.42)						
Log Migration density** (IV non spatial)	0.3499***							
	(4.92)							
Log Migration density** (IV spatial)		0.486***						
		(4.47)						
Log population density(t-1)	0.6467***	0.1588***						
	(3.12)	(2.34)						
Dummy 2018	0.3029***	0.1464***						
	(2.86)	(1.30)						
Dummy pandemic 2020	-0.05***	-0.18***						
	(4.55)	(2.74)						
Spatial rho	0.2988***	0.2988***	0.2988***	0.2988***			0.0188***	0.0991***
lambda					0.0956***	0.0956***	0.0188***	0.0136***
Variance								
sigma2_e	16.53***	16.53***	13.50***	13.62***	18.84***	18.13***	33.56***	32.57***

continue

Variable	SDM IV1	SDM IV2	SAR IV1	SAR IV2	SEM IV1	SEM IV2	SAC IV1	SAC IV2
Direct effects								
Log Formal density(t-1)	0.336***	0.068**	0.300***	0.336**			0.312**	0.154**
	(1.30)	(0.51)	(0.25)	(0.25)			(0.60)	(0.25)
Log Migration density** (IV non spatial)	1.142***		0.881**				0.204**	
	(7.44)		(0.20)				(0.59)	
Log Migration density** (IV spatial)		1.162***		1.0001				0.422
		(5.32)		(0.11)				(0.12)
Log population density(t-1)	0.085**	0.393***	0.529**	0.559***			0.129*	0.673*
	(0.21)	(0.94)	(0.25)	(0.25)			(0.58)	(0.25)
Dummy_2018	0.251	0.143	1.331	1.281			0.222	0.5282
	(0.11)	(0.07)	(0.25)	(0.25)			(0.59)	(0.25)
Dummy_pandemic 2020	-11.34***	-16.80**	-2.10**	-2.19**			0.0655	0.9204
	(0.38)	(0.07)	(0.25)	(0.25)			(0.56)	(0.25)
Indirect effects								
Log Formal density(t-1)	-0.640**	.0.319**	-0.454**	-0.523***			-2.88**	0.0419
	(0.27)	(0.26)	(0.42)	(0.43)			(2.61)	(0.08)
Log Migration density** (IV non spatial)	1.0150**		1.2246***				1.896**	
	(0.11)		(0.43)				(2.71)	
Log Migration density** (IV spatial)		5.00***		3.011***				1.230***
		(0.11)		(0.13)				(0.01)
Log population density(t-1)	0.10245*	0.1012*	2.3893**	2.4521***			1.2187**	0.1877**
	(0.28)	(0.27)	(0.43)	(0.43)			(2.68)	(0.08)
Dummy_2018	0.5643*	0.451**	2.136	2.030			2.071**	0.149
	(0.27)	(0.25)	(0.43)	(0.43)			(2.71)	(0.08)
Dummy_pandemic 2020	-0.758***	-0.572***	-0.276**	-3.459**			0.212*	0.258
	(0.28)	(0.27)	(0.43)	(0.43)			(2.34)	(0.08)
Total effects								
Log Formal density(t-1)	-0.303***	-0.251***	-0.154***	-0.186***			-2.568*	0.1962**
	(11.49)	(6.11)	(5.57)	(6.83)			(2.48)	(3.21)
Log Migration density** (IV non spatial)	2.192***		2.105***				2.1*	
	(4.69)		(1.49)				(2.56)	
Log Migration density** (IV spatial)		6.162***		3.01*				1.652*
		(5.48)		(0.43)				(0.45)
Log population density(t-1)	0.187*	0.494**	2.91***	3.01***			1.3485*	0.861**
	(0.40)	(7.56)	(7.00)	(7.13)			(2.53)	(3.25)
Dummy_2018	0.815	0.594*	3.46	3.312			2.294*	0.6777
	(1.04)	(2.13)	(6.73)	(6.73)			(2.55)	(3.21)
Dummy_pandemic 2020	-12.107***	-17.372**	-2.380**	-5.652**			0.2775	1.178
	(5.54)	(2.53)	(6.64)	(6.99)			(2.23)	(3.24)

continue

Variable	SDM IV1	SDM IV2	SAR IV1	SAR IV2	SEM IV1	SEM IV2	SAC IV1	SAC IV2
Lr-test	6325.5***	4852.5***	12450***	12456***	15679***	13897***	10298.3	10245.6***
SDM Vs SAR χ-test	202.45***	198_304***						
SDM Vs SEM χ-test	58.24**	62.37**						
SDM Vs SAC AIC	8795.1	8215.5	10168.1	10042.1	19676.0	9413.2	19676.0	12912.1
SDM Vs SAC BIC	7697.0	8467.3	11418.8	8987.4	15304.3	9191.6	15304.3	11982.2
Hausman test: IV non Spatial migrants 2005 vs intial model	12.5		8.2		0.19		0.54	
F First Stage IV1	104.6		104.6		104.6		104.6	
Montiel Pflueger Robust IV1	88.23		88.23		88.23		88.23	
IV Spatial migrants 2005 vs intial model		134.52**		100.3*		20.1		12.3
F First Stage IV2		197.61		197.61		197.61		197.61
Montiel Pflueger Robust IV2		96.23		96.23		96.23		96.23
Time effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.6	0.5	0.32	0.34	0.36	0.54	0.53	0.7
R2 adj	0.54	0.48	0.3	0.32	0.34	0.5	0.5	0.6
N	12.881	12.881	12.881	12.881	12.881	12.881	12.881	12.881

Note: “IV1” refers to take the first stage non-instrumental regression. While “IV2” refers to the first stage spatial instrumental regression. +The control variables included the lags of the Logarithm of the informal density of establishments one period behind t-1. The same for the native population density; the Log distance to the CBD (Unicentro) and two dummies (one that takes the value after 2018 when the highest peak of migration and the last one, a dummy that denotes Covid-19 after 2020). Errors are clustered by neighborhood. The Lr-Test of column 2 is performed contrasting SAM model against nesting models SAR-SEM-SAC. The Chi-2 test allows to compare the SDM model against a SAR or SEM model. While the AIC-BIC criteria allow to rule out the SAC model in favor of SDM model. The variable distance to the CBD is also included.

Table 7A. Part 1.Instrumental Variables Estimation for immigration and formal firms association by size.

Main	Ln Big firms	Ln Median firms	Ln Micro firms	Ln Small firms
Local effects	0.0118***	0.0190**	0.2159***	0.1357***
	(0.00220)	(0.00612)	(0.0792)	(0.0197)
Wx: Global effects	0.0737***	0.00501***	0.1923***	0.213***
	(0.0000221)	(0.0000532)	(0.000689)	(0.0138)
Direct effects	0.0164***	0.00134	0.646**	0.5751***
	(0.00220)	(0.00301)	(0.106)	(0.0138)
Indirect effects	-0.000207	-0.000745	0.4195***	0.1294***
	(0.00192)	(0.00271)	(0.0967)	(0.0125)
Total effects	0.0162***	0.000599	1.0655***	0.7045***
	(0.00104)	(0.000473)	(0.0162)	(0.00352)
Time effects	Yes	Yes	Yes	Yes
All controls+	Yes	Yes	Yes	Yes
Spatial rho (W*Formal density)	0.299***	0.299***	0.299***	0.299***
	(0.0000997)	(0.000103)	(0.0000982)	(0.0000997)
Variance				
sigma2_e	0.522***	0.0673***	87.45***	5.421***
	(0.00624)	(0.000804)	(-1,045)	(0.0648)
N	12.881	12.881	12.881	12.881
R-sq	0.58	0.43	0.36	0.43
adj. R-sq	0.55	0.41	0.32	0.41
F-First Stage IV2	197.61	197.61	197.61	197.61
Montiel Pflueger Robust IV2	96.23	96.23	96.23	96.23
SDM Vs SAR $\chi2 - test$	9793.07***	22166.69***	50277.27***	38792.20***
SDM Vs SEM $\chi2 - test$	160.79***	60.43***	102.64***	164.96***
SDM Vs SAC AIC	4948.7	23844.3	68930.4	30821.0
SDM Vs SAC BIC	4752.3	24040.6	68990.8	30881.5
SEM AIC	31168.3	70098.9	95808.4	56733.3
SEM BIC	31107.9	81059.3	96004.7	56929.6
SAR AIC	2948.9	41811.6	122279.4	80952.1
SAR BIC	3099.9	41962.6	122430.5	81103.1
SAC AIC	28303.8	84040.8	159608.7	105194.8
SAC BIC	28462.4	82460.2	159767.3	105353.4

Note: Errors are clustered by neighborhood. Big firms refer to those establishments with more than 200 employees. Median firms are between 51-200 employees. Small firms with 11-50 employees. Micro firms are those that hire less than 10 employees. The Chi-2 test allows to compare the SDM model against a SAR or SEM model. While the AIC-BIC criteria allow to rule out the SAC model in favor of SDM model.

Table 7A. Part 2. Instrumental Variables Estimation for immigration and informal firms association by size.

Main	Ln Big firms	Ln Median firms	Ln Micro firms	Ln Small firms
Local effects	0.0143***	0.0267***	0.242**	0.1463**
	(0.00198)	(0.00564)	(0.0794)	(0.0166)
Wx: Global effects	-0.000246***	-0.00391***	1.0802***	0.9160***
	(0.0000223)	(0.0000646)	(0.000813)	(0.000170)
Direct effects	-0.000656	0.0119**	2.270***	0.9584***
	(0.00421)	(0.00447)	(0.0331)	(0.00431)
Indirect effects	-0.00119	-0.00156	8.00236***	1.00424***
	(0.00381)	(0.00401)	(0.0286)	(0.00675)
Total effects	-0.00185***	0.0103***	10.272***	1.9626***
	(0.000518)	(0.00103)	(0.0137)	(0.00753)
Time effects	Yes	Yes	Yes	Yes
All controls+	Yes	Yes	Yes	Yes
Spatial rho (W*Formal density)	0.299***	0.299***	0.299***	0.276***
	(0.000101)	(0.000104)	(0.0000993)	(0.000123)
Variance				
sigma2_e	0.0562***	0.464***	91.86***	4.031***
	(0.000672)	(0.00554)	(0.01098)	(0.0482)
N	12.881	12.881	12.881	12.881
R-sq	0.43	0.65	0.457	0.320
adj. R-sq	0.30	0.58	0.45	0.3
F-First Stage IV2	197.61	197.61	197.61	197.61
Montiel Pflueger Robust IV2 Test	96.23	96.23	96.23	96.23
SDM Vs SAR χ2-test	10657.63***	22229.45***	42179.67***	40352.21***
SDM Vs SEM χ2-test	472.14***	208.57***	282.07***	266.06***
SDM Vs SAC AIC	641.7	22182.9	69769.4	29244.0
SDM Vs SAC BIC	792.8	22379.2	69829.8	29304.4
SEM AIC	31967.3	22239.1	96499.8	53409.7
SEM BIC	31906.9	22778.7	96696.1	53606.0
SAR AIC	7461.1	39205.7	119904.1	78569.1
SAR BIC	7264.8	39356.7	120055.1	78720.1
SAC AIC	20276.5	75690.3	153567.7	112041.9
SAC BIC	20435.0	75848.9	153726.2	112200.5

Note: Errors are clustered by neighborhood. Big firms refer to those establishments with more than 200 employees. Median firm are between 51-200 employees. Small firms with 11-50 employees. Micro firms are those that hire less than 10 employees. The Chi-2 test allows to compare the SDM model against a SAR or SEM model. While the AIC-BIC criteria allow to rule out the SAC model in favor of SDM model. Last part of the table appears the AIC-BIC criteria values for the other models SEM-SAR-SAC to show that SDM is the best model.

Table 7A. Part 3. Instrumental Variables Estimation for immigration and formal firms association by sector.

Main	Ln Commerce	Ln Industry	Ln High services	Ln Low services	Ln contruction SAC
Local effects	0.650***	0.355*	0.0536	0.936**	0.00357***
	(0.0183)	(0.000316)	(0.0534)	(0.000249)	(0.000508)
Wx: Global effects	0.191***	0.164***	0.473***	0.121***	
	(0.000160)	(0.000135)	(0.000465)	(0.000102)	
Direct effects	0.0663***	0.0568***	0.171***	0.408***	0.00367***
	(0.00600)	(0.00883)	(0.0408)	(0.00520)	(0.000496)
Indirect effects	0.7100***	0.00149	0.00686	0.6310***	0.0917***
	(0.00532)	(0.00787)	(0.0367)	(0.00456)	(0.00827)
total effects	0.7763***	0.0583***	0.178***	1.039***	0.0954***
	(0.00347)	(0.00272)	(0.00969)	(0.00203)	(0.00813)
Time effects	Yes	Yes	Yes	Yes	Yes
All controls	Yes	Yes	Yes	Yes	Yes
Spatial					
rho (W*Formal density)	0.287***	0.299***	0.299***	0.299***	0.2619***
	(0.000121)	(0.0000988)	(0.0000975)	(0.000100)	(0.0000415)
Variance					
sigma2_e	4.679***	3.339***	39.75***	1.915***	0.242***
	(0.0560)	(0.0399)	(0.475)	(0.0229)	(0.00265)
lambda					0.108***
					(0.0000225)
N	12.881	12.881	12.881	12.881	12.881
R-sq	0.31	0.32	0.31	0.31	0.272
adj. R-sq	0.29	0.29	0.29	0.3	0.252
F-First Stage IV2	197.61	197.61	197.61	197.61	197.61
Montiel Pflueger Robust IV2	96.23	96.23	96.23	96.23	96.23
SDM Vs SAR χ2-test	38295.52***	41416.93***	32878.04***	41780.42***	32878.04***
SDM Vs SEM χ2-test	370.56***	658.04***	95.85***	108.55***	95.85***
SDM Vs SAC AIC	29682.7	24317.5	57498.3	10585.3	49442.4
SDM Vs SAC BIC	29743.1	24377.9	57558.7	10743.8	49638.7
SEM AIC	55033.8	49924.7	84730.3	17552.2	24912.4
SEM BIC	55230.1	50121.0	84926.7	17612.6	24972.9
SAR AIC	79248.5	73977.3	101699.0	42108.4	68787.1
SAR BIC	79399.5	74128.3	101850.0	42304.7	68938.1
SAC AIC	113225.4	106379.4	116219.2	61533.3	18174.4
SAC BIC	113383.9	106537.9	116377.8	61684.3	18333.0

Note: Errors are clustered by neighborhood. The Chi-2 test allows to compare the SDM model against a SAR or SEM model. While the AIC-BIC criteria allow to rule out the SAC model in favor of SDM model. Last part of the table appears the AIC-BIC criteria values for the models SEM-SAR-SAC to show that these values are higher than the one obtained for the SDM model. However, the best model chosen for the building sector was a SAC Model. This is shaded by gray color at the right-down part of the table, because de AIC-BIC criteria were the lowest. The chi-2 chooses the SDM against SAR-SEM models, but the SDM is not preferred when compared to the SAC Model.

Table 7A. Part 4. Instrumental Variables Estimation for immigration and informal firms association by sector.

Main	SDM	SDM	SEM	SDM	SDM
	Ln Commerce	Ln Industry	Ln high services	Ln Low Services	Ln Construction
Local effects	0.654***	0.143***	0.0381***	0.8114***	0.258***
	(0.0276)	(0.0140)	(0.00908)	(0.0112)	(0.00592)
Global effects	0.9475***	0.800***		0.916***	0.528***
	(0.0276)	(0.0140)		(0.0115)	(0.0606)
Direct effects	0.376***	0.112*		3.332***	0.129**
	(0.0557)	(0.0386)		(0.00746)	(0.00804)
Indirect effects	2.0123***	0.768**		3.0151***	0.179**
	(0.00503)	(0.0349)		(0.00658)	(0.00723)
Total effects	2.3883***	0.88**		6.33351***	0.308***
	(0.00704)	(0.00431)		(0.00203)	0.00127)
Time effects	Yes	Yes	Yes	Yes	Yes
All controls+	Yes	Yes	Yes	Yes	Yes
Spatial					
rho	0.299***	0.299***		0.299***	0.299***
	(0.0000989)	(0.0000983)		(0.0001000)	(0.0000993)
lambda			0.1956***		
			(0.0000224)		
Variance					
sigma2_e	1.108***	2.864***	1.651***	1.840***	0.511***
	(0.132)	(0.0342)	(0.0197)	(0.0220)	(0.00611)
N	12.881	12.881	12.881	12.881	12.881
R-sq	0.310	0.320	0.225	0.288	0.217
adj. R-sq	0.29	0.31	0.21	0.25	0.19
Lr-test					
F-First Stage IV2	197.61	197.61	197.61	197.61	197.61
Montiel Pflueger Robust IV2	96.23	96.23	96.23	96.23	96.23
SDM Vs SAR χ2-test	30108.96***	38276.98***	50.39	34364.50***	23316.12
SDM Vs SEM χ2-test	430.94***	976.29	17.90	193.70***	290.64
SEM Vs SARχ2-test			600.93***		
SDM Vs SAC AIC	40224.7	21484.6	73885.2	14335.4	20045.3
SDM Vs SAC BIC	40285.1	21545.0	74081.5	14494.0	19840.9
SEM Vs SAC AIC			46386.2		
SEM Vs SAC BIC			46446.6		
SEM AIC	66777.3	47764.6		15768.0	23559.2
SEM BIC	66973.6	47960.9		15828.4	23755.6
SAR AIC	83112.7	67202.5	86132.4	61029.3	37365.0
SAR BIC	83263.7	67353.5	86283.4	61180.3	37516.0
SAC AIC	102391.8	93349.3	91925.1	41550.2	54820.9
SAC BIC	102550.4	93507.9	92083.6	41746.5	54979.4

Note: Errors are clustered by neighborhood. Big firms refer to those which hire more than 200 employees. Median firm are those between 51-200 employees. Small firms are those that hire 11-50 employees. Micro firms are those that hire less than 10 employees. The Chi-2 test compares the SDM model against a SAR or SEM model. While the AIC-BIC criteria allow to rule out the SAC model in favor of SDM model.

Figure 2A. Significance test for non-stationarity

Geographically Weighted Regression

Significance Test for Bandwidth

Observed	P-Value
55000	0.000

Significance Tests for Non-Stationarity

Variable	Si	P-Value
Constant	2.8192	0.900
Log Informal density (t-1)	0.1286	0.002
Log population density (t-1)	0.1420	0.006
Log Migration density (t-1)	0.3200	0.000
Log distance to CBD	0.1924	0.001
Dummy_2018	0.0000	1.000
Dummy pandemic 2020	0.0000	1.000

Figure 3A. K-means number of optimal clusters

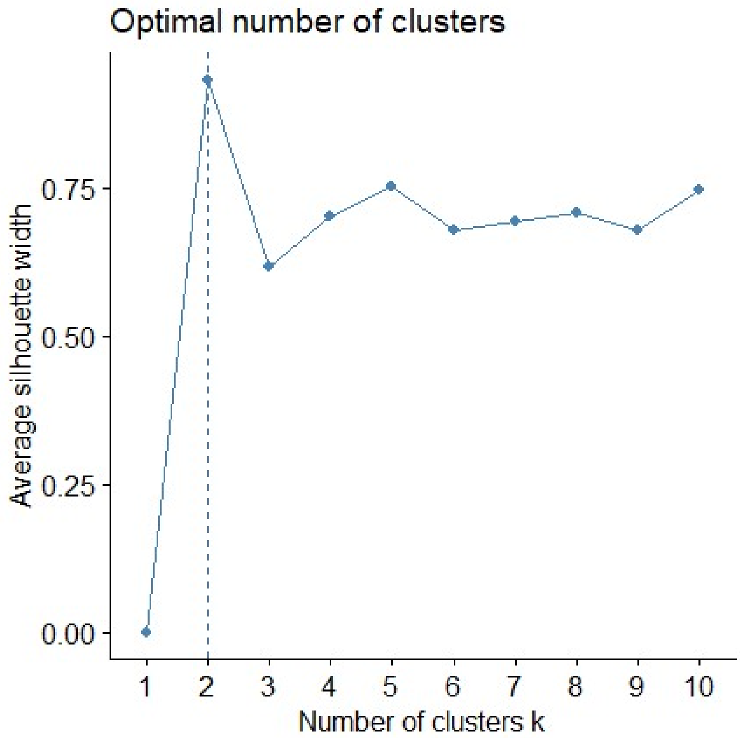
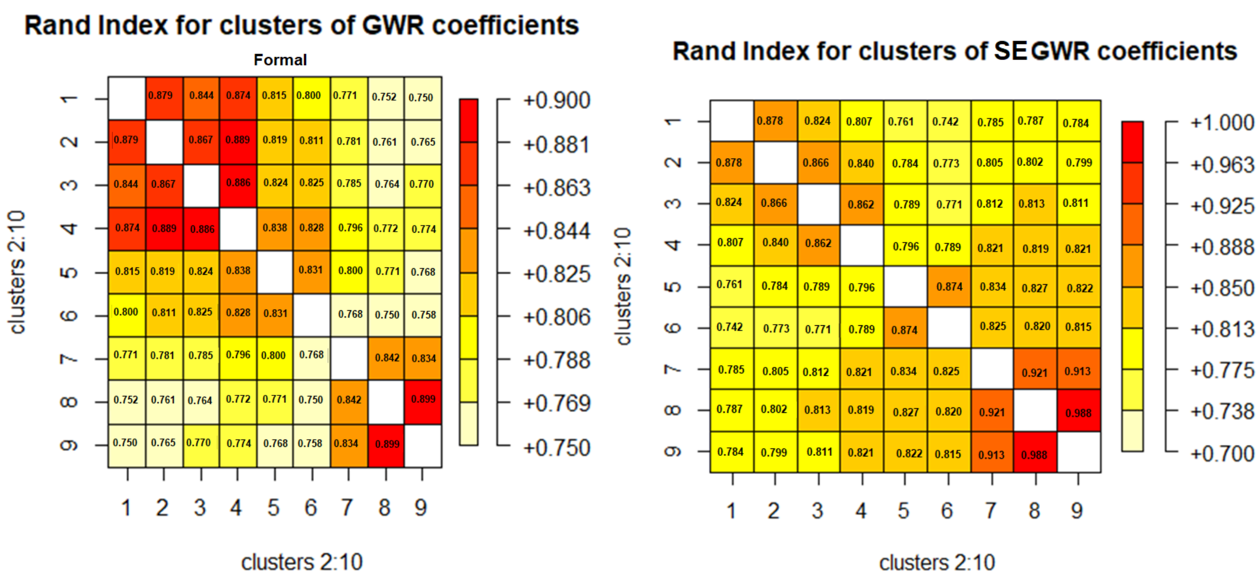


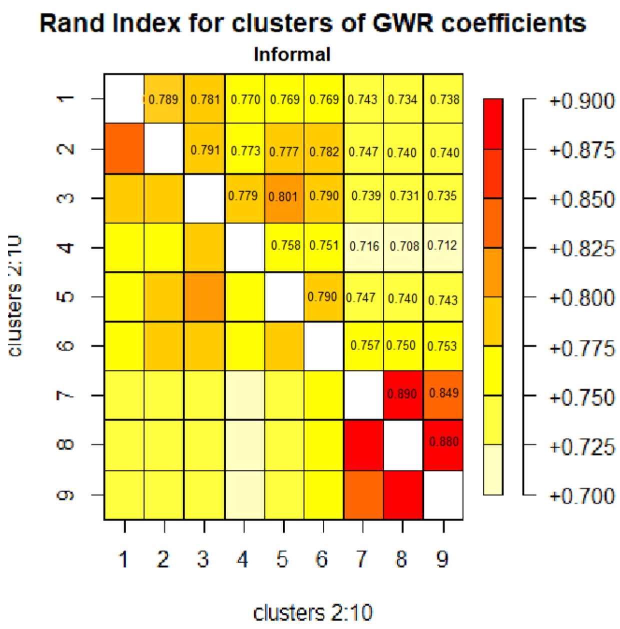
Figure 4A. Rand Index for GWR coefficients of formal firms



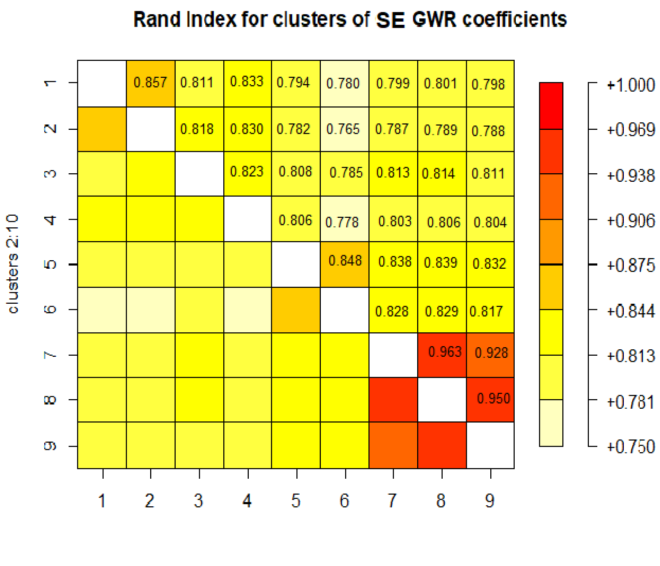
Coefficients

Standar Errors

Figure 4A. Rand Index for GWR coefficients of informal firms



Coefficients



Standar Errors