THE EFFECT OF A LABOR SUPPLY SHOCK ON FACTORS PRODUCTIVITY: THE CASE OF A VENEZUELAN MIGRATION IN COLOMBIA

Diego Armando Martínez Moya[1] and Juan Carlos Muñoz Mora[2]

[1]Thesis to apply for a Master's degree in Economics, School of Economics and Finance, EAFIT University

[2]Advisor, School of Economics and Finance, EAFIT University

April 2020

Abstract

This document analyses the potential effects of a labor supply shock (Venezuelan immigration) on factors productivity in Colombia. Using aggregate data on employment and output in the Colombian departments, we find that under the 2014-2018 time period, Venezuelan immigration has significant effects on multi-factor productivity and other product-related variables in the Colombian departments. The empirical strategy used shows that the mechanism of transmission of the effect is the skill bias of Venezuelan immigrants.
Keywords: Immigration, Productivity, labor market.

1 *INTRODUCTION*

Colombia is facing the largest international migration in its history (Bahar et al., 2018). According to Migración Colombia (2019), in 2018 a total of 1.072.432 Venezuelans were living in Colombia, 166% more than 2017. In the history of Venezuelan migration to Colombia, there are two main migration waves. First, during the first decade of 21st century some Venezuelans with high skills and resources moved into Colombia to find new opportunities for investment or access to labor markets with high skill jobs. Second a wave was characterized by massive immigration catalogued as a migration crisis (FIP, 2018). The latter is considered a refugee crisis (World Bank Group, 2018), and it is driven by the collapse of production and the international price of oil which translate into a significant reduction in waves for Venezuelan workers (Universidad del Rosario, 2018).

This new migration flow brings challenges for Colombia's economy. On the one hand, the first challenge is aimed at the area of service provision required by these immigrants, Colombian Government should cover the basic needs of migrants such as health, housing, education, and security, among others, it implies an increase in public expenditure (Universidad del Rosario, 2018). On the other hand, it is necessary to introduce active social integration programs to reduce the potential impact on things like civil security and labor markets by allowing Venezuelan migrants to properly integrate into Colombian society (FIP (2018); World Bank Group (2018)), paying special attention to the informal sector in the labour market (Calderón-Mejía and Ibáñez, 2016).

Nonetheless, this migratory flow might be an opportunity for the Colombian economy. New Venezuelan immigrants are willing to accept salaries well below the minimum wage established by Colombian law, this seems to be being taken advantage of by employers with the aim of lowering their labour costs, thus tending to increase their workforce (Orrenius and Zavodny, 2012). In this way, this migratory flow is an increase in the supply of labor within the labor market in Colombia, modifying the factor productivity that is characterized for a poor historical performance (ANIF, 2017). According to ANIF (2017), the lack of multifactorial productivity growth is due to high labor costs of up to 50%, high transport costs due to poor road infrastructure and high energy costs with prices that are not competitive at the international level. Therefore, the possible effect of Venezuelan migration on Colombia may shed light on what is happening to departmental productivity in Colombia.

Although there is evidence on the relationship between productivity and immigration in developed countries (Peri, 2012; Lewis, 2011; Hunt and Gauthier-Loiselle, 2010; di Giovanni et al., 2015), very little is known on the direction and magnitude of such relation in developing countries where labour markets are not flexible (Hanson, 2010).

This paper aims at identifying the effect of Venezuelan immigration on total factor productivity in Colombia between 2014 and 2018. After reviewing the main stylized facts of the migration and the factor productivity in Colombia, we built a set of hypothesis based on theoretical framework built by Peri (2012) which explains the main mechanism through which an increase in the labour supply generated by migratory flows can positively or negatively affect multifactor productivity depending on the share of hours worked by high and less educated workers and taking into account the effect of such flows on a productivity intensity index. Then, we estimated a reduced model through a rich data set on migration and productivity at the department level: Gross Departmental Product (GDP) from Departamento Administrativo Nacional de Estadística (DANE), the Gran Encuesta Integrada de Hogartes (GEIH) and finally the Gross Domestic Product by expenditure approach (DANE). The hypotheses that will be raised seek to prove that Venezuelan immigration has a positive effect on hours worked in Colombia; however, that such immigration generates a significant negative effect on productivity since a index of intensity in productivity obtains the contrary effect; that effect can be obtained if such a migration flow increases the share of hours worked by less educated workers and if high educated workers become less productive than less educated workers. According to Peri (2012) the behaviour of these variables are key to assimilating the mechanism that leads to an understanding of the relationship between immigration and productivity.

The results show that Venezuelan immigration has a significant negative effect on departmental productivity in Colombia, reducing its growth by a 48.46%, in part due to the negative effect of migration flow on the skill bias in productivity, meaning that highly skilled workers are becoming less productive than less skilled workers, thus generating that when employment undergoes a 1% increase due to the migration shock, the skill intensity index will increase by 2,084%. However, according to the literature that will be presented in the following sections, this migratory flow has a significant positive effect on total hours worked.

This research contributes to the literature in three ways. First, to the best our knowledge, our results are the first to try to understand the relationship between migration and total factor productivity (TFP) in a developing country. Second, our results uncover an aggregate level of the negative on migration in developing countries. Third, it gives new empirical evidence for policy makers on the potential intervention to Venezuelan immigration has a direct effect on employment factors like hours worked in the Colombian departments.

The paper is structured as follows: section 2 describes some stylized facts and context of Venezuelan migration; section 3 specifies the theoretical framework; section 4 present and discusses the estimations strategy and dataset; section 5 display the estimation results and some robustness checks and finally section 6 concludes the research.

2 *VENEZUELAN MIGRATION IN COLOMBIA: A LABOR SUPPLY SHOCK*

According to Migración Colombia (2019), Venezuelan migration in Colombian territory has had an exponential growth from 2014 to 2018. By 2014, 291.539 Venezuelans entered Colombia, while by 2018 it was a total of 1.359.815. In general, this immigration generates an increase in the availability of the labor factor, that is, it increases the supply of labor as their labour participation rate is around 75%, which has a cheaper cost due to the socioeconomic conditions under which these people enter Colombian territory. This, without a doubt, can be taken advantage of by employers to reduce the cost of their labor. Most of these immigrants have a high school degree, and are covering the informal sector in the Colombian labor market, which represents a large part of this market. Therefore, the effects that this immigration may have are imminent.

According to official statistics (GEIH), dur-

Table 1: Basic Characteristics of Venezuelan migrants and natives

2014-2018				
Venezuelan Migrants	Native	5 Y	12 M	Born
1. General Characteristics				
1.1 Sex				
Female (%)	51,48	47,51	48,74	49,77
1.2 Age -years (%)				
15 – 25	25,86	29,42	36,54	40,42
26 – 35	22,02	31,99	31,73	35,05
36 – 45	18,76	20,36	17,32	15,98
46 – 55	17,15	11,23	8,57	6,02
56 - 70	16,21	7,00	5,84	2,53
1.3 Marital (%)				
Household Head	40,53	38,23	26,75	29,11
Married	21,87	13,13	13,61	15,28
lives with spouse	95,21	93,20	84,47	89,08
2. Education				
Literacy Rate	95,87	96,99	97,78	98,57
Enrolled in school	13,64	7,10	3,01	6,07
Avg. years of schooling	9,71	9,24	9,62	10,43
2.1 Highest Educational Level Obtained				
None (%)	4,15	2,94	2,25	1,39
Preschool (%)	0,01	0,01	0,00	0,00
Basic Primary (1-5) (%)	24,14	20,86	16,58	9,83
Basic Secondary (6-9) (%)	17,56	24,92	24,22	22,21
High School (10 -11) (%)	28,41	31,87	34,29	34,98
Superior or University (%)	25,72	19,40	22,65	31,60
3. Health and Social Security				
3.1 Access to Social Security in Health				
Yes (%)	92,42	60,48	25,47	30,54
3.2 Type of Affiliation				
Contributive (%)	49,83	38,28	34,67	57,15
Special (%)	3,67	0,44	0,17	1,11
Subsidiary (%)	46,44	61,21	65,09	41,67
4. Labor Market				
Participation Rate (%)	65,10	72,65	74,13	70,15
Employment Rate (%)	56,08	65,52	59,25	62,75
Unemployment Rate (%)	9,60	9,81	20,19	18,36
4.1 Pension Contribution				
Yes (%)	37,88	19,71	8,09	15,00
No (%)	60,64	80,21	91,73	84,76
Pensioner (%)	1,49	0,08	0,18	0,24

Source: Elaborated by migration group of EAFIT University with GEIH data ing 2014-2018 it is possible to distinguish between three type of migrants: people who lived in Venezuela 12 months and 5 years ago (short and long term immigrants), and those who where born in Venezuela. Table 1 summarizes basic characteristics of Venezuelan immigrants and Natives (Colombian) between 2014-2018, and provides information stating that the majority of Venezuelan migrants are men although participation is not much greater than that of women, the age range that predominates is 15-25 years, that is, much of the migration is at most young or young adult. In addition, it is important to mention that around 30% of immigrants are heads of household, while if we look at Colombians their rate reaches around 40%, and adding that approximately 15% of these immigrants are married and 89% of them live with their respective spouses; An important condition of these immigrants is that around 30% have a high school degree, and it can be observed for all population groups that the average years of education are related to the degree achieved; on the side of health and social security it is observed that long term immigrants have more access to these services, therefore this indicates that there is a period of time in which such immigrants take to adapt to society, not to mention that around 60% have a subsidized type of affiliation. Finally, this immigration has a labor market participation rate of around 70% and an employment rate of around 60%, but informality prevails since most of these workers do not contribute to pensions.

However, as of 2017, this migratory flow begins to be seen as a migratory crisis, where there was an excessive income of Venezuelans without any type of documentation and willing to work for any salary below the minimum wage in Colombia. Migración Colombia created the Permiso Especial de Permanencia (PEP), with the objective of regularizing the migratory situation of Venezuelans who have expired their permits to stay in Colombia, this is how the Colombian government presents the National Council for Economic and Social Policy (Departamento Nacional de Planeación, 2018), which establishes the public policy guidelines to address this crisis.

Immigrants and natives are not under the same conditions in terms of access to the labour market, immigrants become cheaper labour, which is taken advantage of by the main actors on the demand side (companies), as this lowers their costs, systematically increasing their demand for labour[1]. As an informal worker, those Venezuelan migrants labor receives lower salary, they must work harder and, in some cases, they need to have a second job to get the enough income to survive (FIP, 2018).

The above allows us to begin to observe the mechanism by which this migratory shock can have an effect on productivity at the departmental level, that is, immigration undoubtedly increases the labor supply by increasing the number of people available for the labor market, but how the mechanism can be completed to affect productivity; in this research we will take into account the individual's skill through the highest education degree, therefore, it is thought that these immigrants are less skilled, this leads to a decrease in productivity.

Based on the Gran Encuesta Integrada de Hogares (GEIH), we can see in Figure 1 and 2 the total number of Venezuelan immigrants in each of the departments of Colombia in 2014 and 2018 respectively. By 2018 the 5 departments with the highest number of Venezuelan immigrants was: Bogotá with a total of 266.413 immigrants in 2018 while in 2014 it had a total of 15.409 Venezuelan immigrants; followed by Atlántico, going from 8.995 in 2014 to 150.628 in 2018; Norte de Santander in 2014 went from 20.680 Venezuelan immigrants to 141.629 in 2018; Antioquia went from 3.185 in 2014 to 99.802 Venezuelan immigrants, and finally the department of Guajira, from 7.818 to 97.766 Venezuelan immigrants from 2014 to 2018. In both graphs, the growth of this population can be seen through space, and it can also be perceived that the destination departments with the greatest number of immigrants main-

[1] According to Migración Colombia (2019), in 2016, 55 companies were sanctioned for hiring Venezuelan migrants without any work permit, which implies that those workers were hired without any social security or contract.

Figure 1: Spatial distribution of Venezuelan immigrants 2014

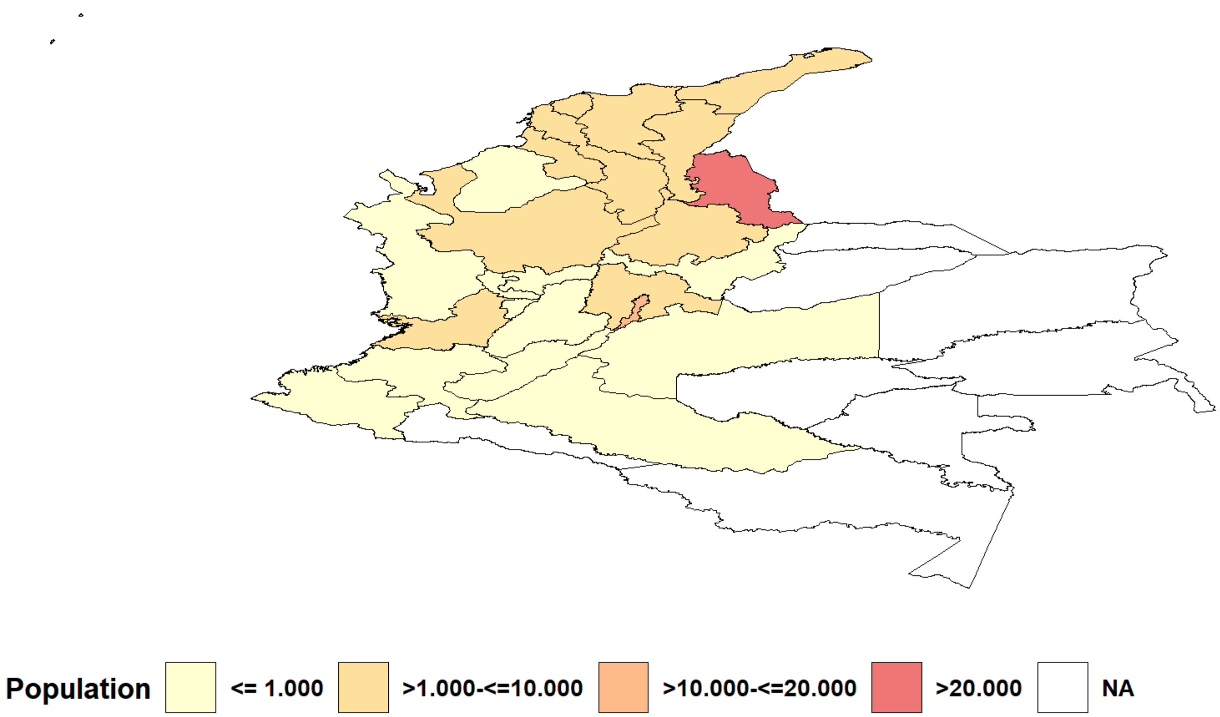
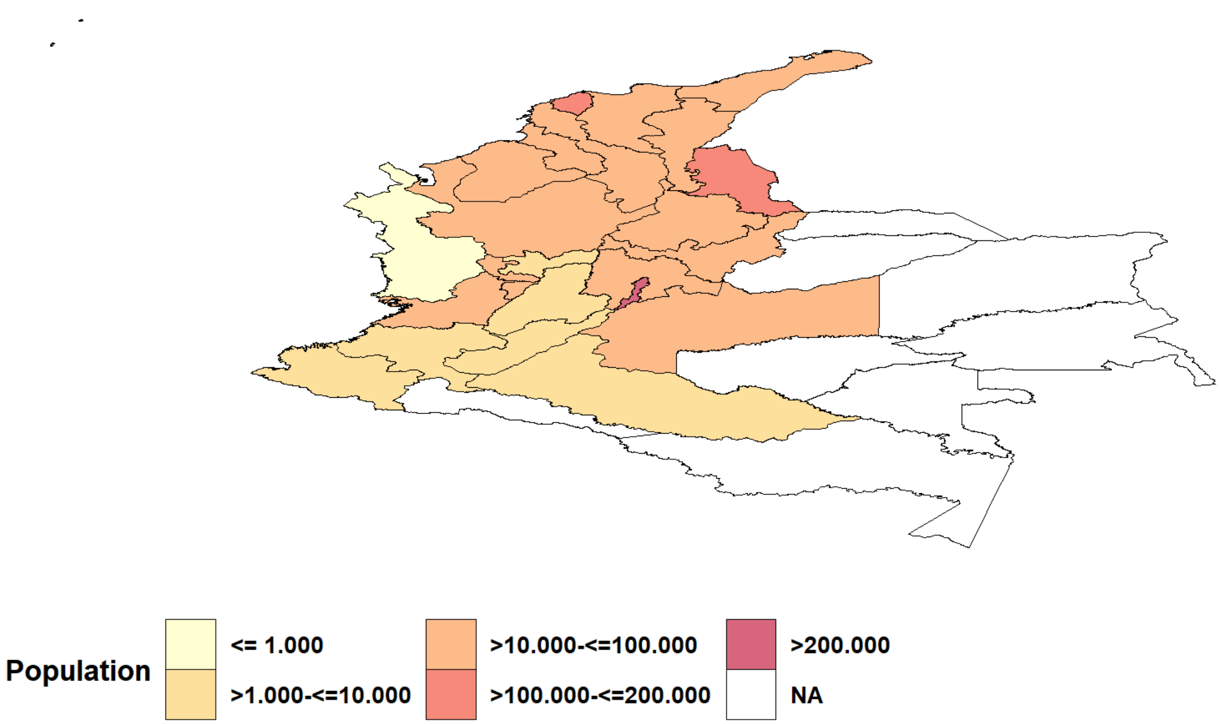


Figure 2: Spatial distribution of Venezuelan immigrants 2018



tain this behavior over time, which leads to the intuition that there is no convergence towards certain types of departments.

Therefore, the stylized facts presented above give free way to affirm that this migratory flow is a strong shock on the labor supply and as a consequence the labor demand as well, the transmission mechanism towards the multifactorial productivity will be presented next. Although, there are rigidities in the Colombian labor market for the access of these immigrants (for example, work permits), the low costs generated by the need of these immigrants to get a job, lead companies to risk hiring them for a lower salary than what is established by law, and making even more effective the effect of immigration on productivity.

The following is a background of seminal work on the effect of immigration on the labour market and productivity.

3 *THEORETICAL FRAMEWORK*

Whereas most scholars have mainly focused on the effect of migration on different labour markets outcomes (e.g. informality, wage, among others) (Altonji and Card (1989); D'Amuri and Peri (2014);Dustmann et al. (2013); Caruso et al. (2019)), relatively less attention has received how these new labor market condition might modify productivity, specially, in developing countries.

According to Barro and Sala.i.Martin (2003) productivity is essential for the growth of an economy in the mid- and long-term, so it is important to analyze how a labor supply shock could modify the capital/labor ratio leading to modification in the productivity factors and the final output. The literature studying the relationship between immigration and productivity mentions different mechanisms by which the effects of migratory inflows on regional or national factor productivity can be studied.

On one the hand, Peri (2012) finds that the transmission mechanism of the migration effect is through an index of skill intensity which contains a parameter that captures the degree of skill-bias of the productivity used in state s and year t and the share of hours worked by high or less educated workers. In summary, if immigration has a negative effect on this index and also a positive effect on the share of high educated workers then this leads to a positive effect on productivity and vice versa. in this seminal work where the author use a production-function that represent the economy of the U.S. states to analyze the impact of immigration on the inputs to production, on productivity, and through these, on income per worker; Peri (2012) present three main findings; (i) the author confirm that there is no evidence that immigrants crowd out employment of (or hours worked by) natives. (ii) find that immigration is significantly associated with total factor productivity growth; and (iii), such efficiency gains are unskilled biased-larger, that is, for less educated workers. These correlations are robust to including several control variables individually (technological adoption, sector composition, openness to international and they are not explained by productivity convergence across states or driven by a few states or decades. the author conjecture that at least part of the positive productivity effects are due to an efficient specialization of immigrants and natives in manual-intensive and communication-intensive tasks, respectively (in which each group has a comparative advantage), resulting in a gain in overall efficiency. Following similar methodology, Lewis (2011) found different results to those obtained by Peri (2012); in particular, he did not find any significant association between low-skill immigration and output per worker, maybe this result because the focus only on the manufacturing sector.

On other hand, immigration may also affect productivity through innovation and possibly through entrepreneurship. A couple of recent studies have focused on immigrants' disproportionate role in patenting and innovation. Hunt and Gauthier-Loiselle (2010) showed that among college graduates, immigrants have much higher patenting rates, which appears to be since foreign college graduates have more education and they specialize in larger proportions in scientific and technological fields. Likewise, Hunt and Gauthier-Loiselle (2010) also found that an increase in foreign college share in a state is associated with an increase in the patenting rate in a state that exceeds what one would expect *mechanically* from the higher patenting rate of immigrants in cross-sectional data, and this leads to an increase in productivity in that state. The author only speculate that this is due to *spillovers*, however, because the patent count data are not broken out by nativity in their panel data. To address this issue, Kerr and Lincoln (2010) linked the names of patent holders to an ethnic names database, which allows them to divide patent counts, not by nativity, but into "Indian", "Chinese" and "Anglo-Saxon" patents. They studied specifically the role of the US high-skill "H1-B" program, and they took advantage of the fact that most H1-B visa holders are Indian and Chinese, making their ethnic groups a reasonable proxy for nativity, they found that areas with more H1-B dependence have moderately higher rates of Anglo-Saxon patenting.

One last mechanism by which immigration may affect average productivity is by increasing product diversity. di Giovanni et al. (2015) estimate the impact of the increase in-product diversity that comes from an increase in the scale of the economy associated with immigration. They found that it has a substantial positive impact on welfare in many immigrant-receiving developed countries.

These mechanisms argue that, under certain conditions, migration could lead to more productivity. However, the theoretical as well the empirical evidence come from developed countries where labor markets has large absorption capacity and are flexible, which means that there are not many barriers to access the labour market, the barriers are more on the cultural side than on the composition of the market itself, a clear example is language. It is still unclear how those mechanisms would work in the developing world.

3.1 *Migration and productivity: a simple model*

This section presents a simply model developed by Peri (2016), which formalize the relationship between migration and productivity. Besides its simplicity and explanation power, this model will allow us to set the main theoretical hypothesis that will be empirically tested for the case of Colombia. In appendix A we provide a detailed explanation of the model.

Peri (2012) assumed that a given state s in year t produces a homogeneous and perfectly marketable production, as follows:

$$Y_{st} = K_{st}^{\alpha}[X_{st}A_{st}\theta(h_{st})]^{1-\alpha} \tag{1}$$

Where Y_{st} is the total production of numeraire goods; K_{st} is the aggregate physical capital; X_{st} is the aggregate hours worked; A_{st} captures the total factor productivity (this is our goal). $\theta(h_{st})$ is a CES function production function, where low (L_{st}) and high (H_{st}) educated workers combine their labor inputs in a with a replacement elasticity of $\sigma > 0$:

$$\theta(h_{st}) = [(\beta_{st}h_{st})^{\frac{\sigma-1}{\sigma}} + ((1-\beta_{st})(1-h_{st}))^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma-1}{\sigma}} \tag{2}$$

Where $h_{st} = H_{st}/X_{st}$s is the share of total hours worked (X_{st}) offered by the most educated workers H_{st}; and, $(1-h_{st}) = L_{st}/X_{st}$ is the share of total hours worked (X_{st}) offered by the less educated workers L_{st}. Thus, $H_{st} + L_{st} = X_{st}$. And, β_{st} captures the degree of skill bias of the productivity used in region s and year t [2].

Now, we transform in equation 1 in per capita terms, dividing by the total employment in the state s and year t (N_{st}):

[2] So, a value close to one implies that highly educated workers are much more productive than less educated ones, and an increase in β implies that the highly educated workers are becoming more productive relative to less educated workers

$$y_{st} = \left[\frac{K_{st}}{Y_{st}}\right]^{\frac{\alpha}{1-\alpha}} [x_{st} A_{st} \theta(h_{st})] \quad (3)$$

where K_{st}/Y_{st} is the product-capital relationship, and $x_{st} = X_{st}/N_{st}$ average hours worked per person. In order to obtain the growth rate we apply the logarithm both sides and express in term of total products, that is:

$$\hat{Y_{st}} = \hat{N_{st}} + \frac{\alpha}{1-\alpha}\left[\frac{\hat{K_{st}}}{Y_{st}}\right] + \hat{A_{st}} + \hat{x_{st}} + \hat{\theta_{st}} \quad (4)$$

The classic approach to economic growth indicates that output growth is due only to increases in productivity while all other inputs remain constant (Barro and Sala.i.Martin, 2003). But in equation 4, total output growths due to an increase of total employment (N_{st}) as well as other factors such as capital share, factor productivity, average hours worked and a productivity-weighted skill intensity index.

Thus, an exogenous shock of migration will affect the economy through different channels. First, a direct positive channel with the total output through a mechanical increase of the total employment (N_{st}). Second, indirect channels through the modification of other factors such as capital share ($\frac{\hat{K_{st}}}{Y_{st}}$), factor productivity ($\hat{A_{st}}$), average hours worked ($\hat{x_{st}}$) and a productivity-weighted skill intensity index $\hat{\theta_{st}}$. In sum, the total effect of an exogenous shock of migration on the total output growth will be the compound effect of both channels.

Therefore, Peri (2012) suggests that giving these relationships, we can express any productive factors ($\frac{\hat{K_{st}}}{Y_{st}}, \hat{A_{st}}, \hat{x_{st}}, \hat{\theta_{st}}$) as function of the changes on the total employment due to an exogenous shock of migration, as follows:

$$\hat{b_{st}} = d_t + d_s + \eta^b \frac{\Delta N^f_{st}}{N_{st}} + \varepsilon_{st} \quad (5)$$

where $\hat{b_{st}} \in [\hat{N_{st}}, \frac{\hat{K_{st}}}{Y_{st}}, \hat{A_{st}}, \hat{x_{st}}, \hat{\theta_{st}}]$, $\frac{\Delta N^f_{st}}{N_{st}}$ is the percentage change in employment due to an exogenous shock of immigrants (N^f_{st}), d_t is a fixed effect of the years, d_s is a fixed effect of the state and ε_{st} is the random shock of mean zero and constant variance. Moreover, η^b is the migration elasticity of a given productive factor, which will be the main parameter of interest.

Given the above equation and taking into account the theoretical framework highlighting that Peri (2012) states that there is a positive and significant relationship between migration and growth of total hours worked ($\hat{x_{st}}$), the following hypothesis can be tested:

Hypothesis 1: The effect of an increase in employment due to immigration generates a positive and significant effect on total average hours worked.

Moreover, factor productivity ($\hat{A_{st}}$) and a productivity-weighted skill intensity index ($\hat{\theta_{st}}$) are not observed. However, we know that the average hourly wage of the low (W^L_{st}) and high (W^H_{st}) educated workers must be equal to their respective marginal productivity (i.e. $W = \frac{\partial Y_{st}}{\partial h}$ where $h \in [L, H]$). So, we can express the productivity ratio as follows:

$$\frac{W^H_{st}}{W^L_{st}} = (\frac{\beta_{st}}{1-\beta st})^{\frac{\sigma-1}{\sigma}} (\frac{h_{st}}{1-h_{st}})^{-\frac{1}{\sigma}} \quad (6)$$

Therefore, we can obtain a closed expressions of β_{st} in terms of observable variables W^H_{st}, W^L_{st}, given by:

$$\beta_{st} = \frac{(W^H_{st})^{\frac{\sigma}{\sigma-1}} (hst)^{\frac{1}{\sigma-1}}}{(W^L_{st})^{\frac{\sigma}{\sigma-1}} (1-h_{st})^{\frac{1}{\sigma-1}} + (W^H_{st})^{\frac{\sigma}{\sigma-1}} (h_{st})^{\frac{1}{\sigma-1}}} \quad (7)$$

Moreover, by replacing β_{st} in equation 2, we can express the productivity-weighted skill intensity index ($\hat{\theta_{st}}$) as follows:

$$\theta(h_{st}) = \frac{((W^H_{st})(h_{st}) + (W^L_{st})(1-h_{st}))^{\frac{\sigma}{\sigma-1}}}{(W^L_{st})^{\frac{\sigma}{\sigma-1}} (1-h_{st})^{\frac{1}{\sigma-1}} + (W^H_{st})^{\frac{\sigma}{\sigma-1}} (h_{st})^{\frac{1}{\sigma-1}}} \quad (8)$$

Finally, we proceed replace equation 8 into equation 1 and clear factor productivity ($\hat{A_{st}}$):

$$A_{st} = \left(\frac{Y_{st}^{\frac{1}{1-\alpha}} K_{st}^{-\frac{\alpha}{1-\alpha}}}{X_{st}} \right) \times \frac{(W_{st}^{L})^{\frac{\sigma}{\sigma-1}}(1-h_{st})^{\frac{1}{\sigma-1}} + (W_{st}^{H})^{\frac{\sigma}{\sigma-1}}(h_{st})^{\frac{1}{\sigma-1}}}{((W_{st}^{H})(h_{st}) + (W_{st}^{L})(1-hst))^{\frac{\sigma}{\sigma-1}}} \quad (9)$$

Now that TFP ($\hat{A_{st}}$) is an observable variable, equation 5 can be used with $\hat{A_{st}}$ as the dependent variable and remembering that Peri (2012) finds that there is a positive relationship between immigration and productivity, the following hypothesis can be tested:

- *Hypothesis 2: An increase in employment caused by an increase in the migratory flow of Venezuelan employees generates a significant negative effect on productivity in Colombia*

4 Empirical Framework

4.1 Data

We aim at estimating the effect on an exogenous shock of migration in productivity using the case of Colombia. For doing this we built a balanced panel of 24 departments (regions) from 2014 to 2018, combining three main data sources. First, the Gross Departmental Product (Y_{dt}) at constant prices (2015) was obtained from Departamento Administrativo Nacional de Estadística (DANE).

Second, the Gran Encuesta Integrada de Hogares (GEIH) was used to calculate the variables corresponding to employment (N_{dt}), total hours worked X_{dt} and wages W_{dt}(which were deflated with the 2015 Consumer Price Index). Total employment was obtained by adding the Colombians (N_{dt}^{n}) employed (either formally or informally) with the Venezuelan employees (N_{dt}^{m}) over a period of time for each department, therefore $N_{dt} = N_{dt}^{n} + N_{dt}^{m}$. The aggregated hours worked were obtained from the hours worked in the week multiplied by the weeks in a year for any workers in the period of time established for each department; and are added to a group established by their educational level. These educational groups are constructed in such a way that high skill workers are those who have a college or postgraduate degree, and those who make up the low skill group are those who have a barchellor degree, technical and technological degree or no degree at all; in this manner, hours worked by high and low skill workers are represented by $H_{dt} = H_{dt}^{n} + H_{dt}^{m}$ and $L_{dt} = L_{dt}^{n} + L_{dt}^{m}$ respectively.Therefore the aggregated hours worked are represented as follows: $X_{dt} = H_{dt} + L_{dt}$. For wages, the average hourly wage for each department d and in year t is required. This variable is obtained by dividing the monthly salary by the hours worked in a month by each worker[3]. Now, following Medina and Posso (2010) we assume that the substitutability between more and less educated workers $\sigma = 1.47$ and the capital share is $\alpha = 0.3$[4], we can calculate the total factors productivity (A_{dt}) in department d and year t.

Third, the construction of the physical capital stock is a little more cumbersome, since there is no database at departmental level. Therefore, following Peri (2012), we use the output share of each department to calculate the physical capital for each department, which assume a capital-output ratio for all departments. In addition, the proxy for calculating physical capital is gross capital formation, extracted from DANE database that divided from output from the expenditure approach.

Finally, to approach the migration flow we use the migration module of GEIH. Using this information we can build two definition of migration.

[3]It must be clear that once the variables are found, the database is collapsed by year and department, taking into account the individual weight

[4]In this paper, researches use occupations descriptions for Colombia, Brazil, and Mexico, to build computer-use related tasks intensities, in that way can find those coefficients

Table 2: Summary Statistics

	Mean	SD	Min	Max	P25	P50	P75	p99	n
$\hat{y}$	16,35	0,48	14,19	17,52	15,96	16,35	16,67	17,30	120
$\hat{k}$	13,45	0,64	10,98	14,67	12,97	13,45	13,94	14,56	120
$\hat{A}$	7,92	0,58	5,06	9,14	7,51	7,91	8,35	9,08	120
$\hat{h}$	0,19	0,03	0,12	0,31	0,17	0,18	0,20	0,27	120
$\hat{\beta}$	0,59	0,13	0,32	0,85	0,49	0,56	0,69	0,85	120
W_{MB}^H	12492,91	2346,21	6981,54	20439,50	11144,90	12465,84	13524,15	19247,33	120
W_{MB}^L	4008,64	672,25	2395,74	5509,02	3394,64	4038,49	4475,48	5417,64	120
W_M^H	12482,76	2350,77	6952,15	20492,79	11110,24	12455,63	13524,15	19244,54	120
W_M^L	3998,20	673,63	2395,67	5499,60	3375,96	4034,97	4464,60	5408,63	120
N_{Nati}	880538,20	933761,90	75889,44	4976571	320144,10	561187,40	1051080	4880093	120
N_M	14169,28	23934,17	39,02	177555	1289,03	5152,84	15865,13	106797,10	120
N_{MB}	8044,92	17499,55	0	143737	336,67	1985,28	7201,51	69324,18	120

Note:The description and source of the variables can be found in the table 9

First one, defining Migrant as those people who were born in Venezuela, those who 12 months ago lived in Venezuela (short term migrants) or those who 5 years ago lived in Venezuela (long term migrants), this with the objective of capturing the whole migratory flow. Second one, migrant as a person who was born in Venezuela. The reason for using these definitions is that this migratory flow caused many Colombians who lived in Venezuela to return to their places of origin in Colombia, so both Venezuelan and Colombian labor entered, but it should be clarified that we will look mainly at the effects of the second definition of migrant.

Some descriptive statistics of the main variables are presented in table 2, Income per worker($\hat{y}$), capital per worker ($\hat{k}$) and productivity ($\hat{A}$) are presented in their logarithmic scale [5]. Here it should be noted that the participation of total hours worked supplied by less educated workers ($\widehat{1-h}$) is greater than that of the most educated ($\hat{h}$), the former having an average participation of 18%, while the more educated with 82%, although for the maximum participation of the most educated workers for a given department in a given year between 2014-2018 reached 31%. Another interesting fact is that the average monthly salary between 2014-2018 for the 24 departments in the sample is around the minimum wage in Colombia established in 2018, which was around $781.242 Colombian pesos. However, one thing to keep in mind is beta, as a value close to 1 indicates that more educated workers are much more productive than less educated workers, and an increase in this value therefore means that more educated workers are becoming more productive than less educated ones.

[5]The description and source of the variables can be found in table 9

4.2 *Estimation Strategy*

Based on our theoretical framework, our main dependent variable will be the total factor productivity ($\hat{A_{dt}}$) in department d in year t could be written as follows:

$$\hat{A_{dt}} = \delta_0 + \eta^A \frac{\Delta N_{dt}^f}{N_{dt}} + \sum_{d=1}^{24} \vartheta_d + \varepsilon_{dt} \tag{10}$$

where $\frac{\Delta N_{st}^f}{N_{st}}$ is the percentage change in employment due to an exogenous shock of immigrants (N_{st}^f), ϑ_d is the set of regional fixed effects and ε_{st} is the random shock of mean zero and constant variance. And, η^A is the elasticity, which will allow us to interpret the effect of immigration on the TFP. Under Hyphotesis 2, we should expect that $\eta^A < 0$.

Moreover, as we discussed in equation 5, we replicate the same specification 10 to the other productive factors (hours worked $(\hat{x_{st}})$ and productivity-weighted skill intensity index $\hat{\theta_{st}}$) in order to obtain specific migration elasticities for each variable (that is η^x, η^θ). Acording to Hyphotesis 1, we should expect that $\eta^x > 0$.

Equation 10 can be estimated by OLS in terms of δ_0, elasticities for each variable (that is η^x, η^θ), including department fixed effects, accounts for state-specific trends in immigration once the observable forms for skill-bias and multi-factorial productivity given by equations 7 and 9 respectively are found.

However, this estimation could suffer of endogenity. In this specific case, endogenity is subject on the side of simultaneity, since our independent variable is measured as a change in employment driven by a change in migrant employment, this variable is related to the productivity of a specific department in a given period of time, because immigrants can simply head for departments with high productivity growth rates, and this would not allow a proper interpretation between productivity and immigration.

For address this issue Peri (2012) uses an instrumental variable that mixes the immigration rate, a geographic variable that consists of calculating the distance between the main entrance to the United States (the US-Mexico border) and each state, and use a dummy variable per decade; this allows them to capture the fact that the distance from the border has a great effect on predicting the flow of immigrants resulting in a greater presence of Mexicans, which fits the data. the author uses a Bartik (or modified shift-share) instrumental variable methodology [6].

In our case, following Carpio and Wagner (2017) and Morales (2018) we will create a Bartik instrumental variable to overcome the endogeneity problem, since this type of instruments are built with an exogenous input, which ensures that our independent variable is not correlated with the error term of the proposed model. This input will be explained below.

The difference between this methodology and that used by Peri (2012) is that the calculation of the distance in this case is not from the borders but from each of the Venezuelan states to each Colombian department. Below is the methodology for calculating this instrumental variable:

$$IV_{st} = \sum_v \frac{1}{T_{vs}} \delta_v \gamma_t \qquad (11)$$

Where T_{vd} is the travel distance between each Venezuelan state v and each Colombian department d; δ_v is the participation of Colombians in each Venezuelan states v before the Venezuelan crisis (in our case it is a period of time between 2000-2010), this input to construct the variable is the one that allows to assure the exogenity of the independent variable, since this participation is not subject to elements related to the migratory crisis in the period of time studied, in the first decade of the 2000 the socioeconomic conditions in Venezuela were totally foreign to the present ones, and the participation of the Colombians in these Venezuelan villages are not linked with some event of the present crisis; and finally, γ_t is the stock of Venezuelan immigrants in Colombia in year t. The distance is calculated using the travel distance between the center of the main city of each state between Colombia and Venezuela [7]; then, we use the Venezuelan Census of 2011 to derive the pre-crisis share of Colombians that resided in a Venezuelan province (INE, 2011); Therefore an identification hypothesis is that immigration rates depend largely on the constraints posed by physical distances between locations. It should be noted that once the instrumental variable has been calculated, equation 5 is estimated, and this estimation process

[6] A clear explanation of what these instruments are and how to use them with their respective modifications is provided by Goldsmith-Pinkham et al. (2018); Adão et al. (2019); Borusyak et al. (2018); Tabellini (2020)

[7] the mechanism to calculate these distances is through Bing's API with an R code that calculates the distances (the inputs are two shapefiles, the one from Colombia and Venezuela)

Table 3: Effects of Immigration on the Components of productivity

	Migrant		Migrant Born	
	(1)	(2)	(3)	(4)
Dependent Variable	OLS	2SLS	OLS	2SLS
$\hat{A}$	-4.372	-79.71***	-25.80***	-48.46***
	(4.630)	(18.46)	(3.448)	(9.066)
$\hat{x}$	0.254	4.624*	2.349**	3.221*
	(0.581)	(2.245)	(0.760)	(1.357)
$\hat{\theta}$	0.0152	3.430*	0.613	2.084*
	(0.355)	(1.540)	(0.445)	(0.880)
$\hat{h}$	-0.127	-10.58**	-2.241	-6.324**
	(0.795)	(3.547)	(1.319)	(2.068)
$\hat{\beta}$	-1.214	-14.80**	-4.484*	-8.950**
	(1.298)	(4.901)	(1.801)	(2.804)
$\widehat{(1-h)}$	-0.0290	2.491**	0.400	1.488**
	(0.178)	(0.904)	(0.314)	(0.525)
R-squared	0.006		0.112	
First-stage F-test		31.52		73.15
Observations	120	120	120	120

Note: The explanatory variable is immigration as a percentage of initial employment. Each cell is the result of a separate regression. The units of Observations are Colombian departments in year 2014–2018. Each regression includes department fixed effects. The method of estimation is least squares with observations weighted by the employment of the state. Errors in parentheses are heteroskedasticity robust and clustered by departments. The calculated variables use the assumption that $\sigma = 1.47$ and $\alpha = 0.3$. *Significant 5%, **1%, ***0.1%.

is used heteroskedasticity robust errors and clustered by departments to account for the arbitrary correlation of outcomes within these locales.

5 *Estimation Results*

Table 3 reports the baseline results for the different regressions that includes department fixed effects, and reports the heteroskedasticity-robust standard errors clustered by department to account for potential correlation of the residuals these locations. The first row of table 3 decompose the effect of immigration on TFP ($\hat{A_{dt}}$); the second row shows the effect on average hours worked growth $\hat{x_{dt}}$; third one row exhibit the effects on skill intensity index ($\hat{\theta}(h_{dt})$); and the last three rows decompose the effect on the previous variable, thus having the effect on share of the most educated workers $\hat{h_{dt}}$, the skill bias in productivity $\hat{\beta_{dt}}$ and the share of less educated workers $(1 \widehat{- h_{dt}})$.

In addition, in table 3 we check the first-stage F-test of the 2SLS models is estimated with the objective of identifying whether the instrument chosen to carry out the estimation is strong enough to replace the independent variable[8]. Let us remember that what the first-stage estimation does is to regress the independent variable against the instrumental variable, to know its significance and magnitude; it is also taken into account that F-test exceed the nominal value of Wald's test by 5% which is met in this case. These models are a bit consistent with the results presented above, however it should be noted that standard errors grow significantly when we look at the estimates based on the Migrant (remember that this definition takes into account short and long-term migrants and those born in

[8]The first-stage estimates can be found in Appendix B

Venezuela) variable, also the first-stage F-test is smaller (6.30) than that presented by the model in (4) column (73.15)[9]. This confirms that the instrument chosen is a strong instrument for our explanatory variable and pay attention to the variables when they are calculated through the definition of Migrant born is a good choice.

Analyzing the results, the side of the variable of interest, we find that immigration has a negative (high magnitude) and significant effect on productivity ($\hat{A_{dt}}$), the elasticity presented in our results means that with 1% increase in employment due to immigrant employment, departmental productivity growth is reduced by 48.5%, which would lead to the departmental average of said variable being around decrease to 3.8%. However, on the side of total hours worked growth ($\hat{x_{dt}}$), we find a significant and positive effect (3.221), that means when employment increase 1% due to migrant shock total hours worked growth increase about 3.22% . The above accompanied by a positive and significant effect on skill intensity index ($\hat{\theta}(h_{dt})$), given that when employment increases by 1% given a migration shock, skill intensity index growth increase on 2.08%. Therefore, in order to be consistent with the mechanism presented by, when employment increase 1% due to migrant shock the share of the most educated workers $\hat{h_{dt}}$, the skill bias in productivity $\hat{\beta_{dt}}$ and the share of less educated workers $(\widehat{1-h_{dt}})$ decrease in 6.32%, -8.95% and increase 1.48% respectively [10].

It is important to observe the behaviour of the effect that the definition of immigrant has on our dependent variables. In general, the effect that immigration has on the different dependent variables proposed when we analyze the definitions of migration, they maintain the significance and effects in both cases; however, there is a change in the magnitudes of the effects on almost all dependent variables. Nonetheless, although the magnitudes of the effects when we observe the Migrant variable increase, at the same time their standard errors do, indicating that the interpretations become a little more ambiguous.

5.1 *Productivity and Skill bias*

Our baselines estimates indicate that there exist a negative effect on productivity, however, one might think that this effect is driven only by the fact the migration shock was concentrated only in on type of education level. Tables 4 and 5 present the results of the effects of immigration on multi-factor productivity when we take into account some controls variables and make a modification in our σ coefficient.

The first and second rows present the results of the OLS and 2SLS models, for the third and fourth rows (in both tables) the results are presented by modifying the substitutability between more and less educated workers (σ). Column (2) introduces the control of research and development expenditure per worker for each department (R&D), since it is a variable that is associated with both productivity and immigration, and column (3) introduces population change as control too.

Focusing on column (2), it can be seen that the control maintains the significance that immigration has on productivity when we take the result of the basic 2SLS, but reduces the magnitude of the effects. Once σ value increase, the effect keep significance and the magnitude decreases. The same seems to happen when we look at the results shown in column (3). Therefore, for research purposes the inclusion of both variables as controls does not change much effect of immigration on productivity (with an elasticity of -37.38 for R&D and an elasticity of -44.25 for population change). Looking now at the same effects for the skill bias in productivity, the scenario seems to change with respect to the previous one, and that is that when we introduce the R&D control the effect that immigration has on this variable is less significant. However, when we apply the population change control the size of the effect is reduced but the significance remains. Now, when the value of sigma is modified, as opposed

[9] this First-stage F-test presented in table ?? are the results of the model where the explanatory variable is the total factors productivity

[10] All elasticities have the same interpretation as the effect on productivity presented above

Table 4: Estimated Impact of Immigration on Total Factor Productivity ($\hat{A}$)

Dependent variable= $\hat{A}$	(1) Baseline	(2) Controlling by R&D per worker	(3) Controlling by Population Change
OLS	-25.80***	-17.17***	-20.45***
	(3.448)	(3.128)	(5.046)
2SLS	-48.46***	-37.38***	-44.25***
	(9.066)	(8.573)	(10.02)
2SLS: A constructed with $\sigma = \infty$	-45.17***	-34.03***	-41.33***
	(8.343)	(7.784)	(9.053)
2SLS: A constructed with $\sigma = 2$	-47.58***	-36.32***	-43.46***
	(8.735)	(8.221)	(9.591)
Observations	120	120	120

Note: Estimations included department fixed effects. The baseline estimate (row 1) is OLS with TFP constructed using the assumption that $\sigma = 1.47$. In the second row, we use 2SLS with imputed immigrants and deparments distance interacted with share of colombians in Venezuela before de crisis as instruments. In the third row, i calculate the TFP as simply a Solow residual without accounting for the imperfect substitution of the more and the less educated. In the fourth row, total factor productivity is constructed under the assumption that $\sigma = 2$, the elasticity of substitution between the more and the less educated is 2.Observations are Colombian departments in year 2014–2018. The method of estimation is least squares with observations weighted by the employment of the state. Errors in parentheses are heteroskedasticity robust and clustered by departments.*Significant 5%, **1%, ***0.1%.

Table 5: Estimated Impact of Immigration on Skill Bias ($\hat{\beta}$)

Dependent variable=$\hat{\beta}$	(1) Baseline	(2) Controlling by R&D per worker	(3) Controlling by Population Change
OLS	-4.528*	-3.332	-3.632
	(1.865)	(2.216)	(1.924)
2SLS	-9.031**	-7.822*	-8.485**
	(2.828)	(2.985)	(3.047)
2SLS: $\hat{\beta}$ calculated with $\sigma = \infty$	-0.436	-0.154	-0.574
	(0.363)	(0.369)	(0.416)
2SLS: $\hat{\beta}$ calculated with $\sigma = 2$	-3.581**	-3.005*	-3.511*
	(1.206)	(1.251)	(1.354)
Observations	120	120	120

Note: Estimations included department fixed effects. The baseline estimate (row 1) is OLS with TFP constructed using the assumption that $\sigma = 1.47$. In the second row, we use 2SLS with imputed immigrants and deparments distance interacted with share of colombians in Venezuela before de crisis as instruments. In the third row, i calculate the TFP as simply a Solow residual without accounting for the imperfect substitution of the more and the less educated. In the fourth row, total factor productivity is constructed under the assumption that $\sigma = 2$, the elasticity of substitution between the more and the less educated is 2.Observations are Colombian departments in year 2014–2018. The method of estimation is least squares with observations weighted by the employment of the state. Errors in parentheses are heteroskedasticity robust and clustered by departments.*Significant 5%, **1%, ***0.1%.

Table 6: Robustness Checks of the Main Effects of Immigration

Dependent variable	(1) Baseline	(2) Without Border Deparments	(3) Without Larger Deparments
$\hat{A}$	-48.46***	-60.47***	-49.88***
	(9.066)	(14.91)	(10.50)
$\hat{x}$	3.221*	4.540**	2.412
	(1.357)	(1.653)	(1.395)
$\hat{\theta}$	2.084*	2.095	2.021*
	(0.880)	(1.458)	(0.945)
$\hat{\beta}$	-8.950**	-7.806	-9.342**
	(2.804)	(4.644)	(3.106)
$\hat{h}$	-6.324**	-6.562	-7.697***
	(2.068)	(3.357)	(2.236)
$(\widehat{1-h})$	1.488**	1.517	1.856**
	(0.525)	(0.851)	(0.559)
Observations	120	100	105

Note: The explanatory variable is immigration as a percentage of initial employment. Each cell is the result of a separate regression. The units of Observations are Colombian departments in year 2014–2018. Each regression includes department fixed effects. The method of estimation is 2SLS with observations weighted by the employment of the state. Errors in parentheses are heteroskedasticity robust and clustered by departments. Border deparments are: Guajira, Cesár, Boyacá and N.Santander; and the largest deparments are: Bogotá D.C, Antioquia and Valle del cauca. The calculated variables use the assumption that $\sigma = 1.47$ and $\alpha = 0.3$. *Significant 5%, **1%, ***0.1%.

to productivity, the skill bias in productivity ($\hat{\beta}$) suffers significantly, since when σ increases β it becomes less and less significant. It should be mentioned that this behavior makes analytical and mathematical sense, since in the end productivity is calculated with the value of θ, but it is the variable that would be less affected by the change of this parameter. It is in this way that the importance of choosing a coherent and accurate value of the sigma parameter is affirmed.

5.2 *Robustness checks and spatial dependence test*

Complementing the results, some estimates are made in order to identify whether this flow of immigrants changes the effect on our dependent variables if we do not take into account the border departments and the departments with the largest economies, which would give us clues as to whether convergence across departments may bias the estimates if immigrants move to certain departments; in general, results shows that there is no convergence towards the larger departments, however immigration seems to have a different behavior when the border departments are omitted, since when employment increases by 1% due to the migratory shock, the elasticities of variables related to the skill intensity index growth ($\theta(h_{dt})$) lose significance; although productivity growth ($\hat{A_{dt}}$) maintains its negative and significant relationship, decreasing 60.47% when the above occur, but loses explanatory power as standard errors increase.

First, table 6 allows us to identify if the estimates made are tied to the convergence of immigrants to the departments with a larger economy or to the border departments, if these estimates remain significant we ensure that the interpretations are not subject to certain specific departments (this is important since these departments are where a greater concentration of Venezuelan immigrants are found). So, if you look at column (2) the effect on productivity growth ($(\hat{A_{dt}})$) it is obtained that the magnitude of the effect increases (-60.47) and maintains significance, and finally, the skill bias in productivity loses significance (as well as all the variables that compose it, shown in columns 4, 5 and 6) and its elasticity becomes 2,095. Second, when we look at column (3), by not taking into account the longer departments, the relationship and significance remains, however the explanatory power of the instrument is reduced due to the increase of standard errors in each of the models. Therefore, even if the explanatory power of the instrument is diminished, it is concluded that the correlations are not explained by the convergence of productivity between states or driven by a few states; However, the effects appear to have a dependency on the border departments, possibly because there are unobserved cultural and socioeconomic effects that keep immigrants tied to these departments.

Table 7: Moran's I over $\hat{A}$

	Migrant		Migrant Born	
Year	Moran's I	P-value	Moran's I	P-value
2014	1.393	0.164	1.164	0.245
2015	0.586	0.558	0.081	0.936
2016	0.552	0.581	0.400	0.690
2017	2.426	0.015	2.412	0.016
2018	2.345	0.019	2.212	0.027

Finally, table 7 indicates the null hypothesis for Moran's I states that the attribute analyzed is not distributed randomly among the n study areas, and although the significance increases over the years, it can be deduced that there is no existence of spatial auto-correlation, this can be seen for both population groups (Migrant and Migrant Born); to carry out this procedure, the island of San Andrés and Providence were excluded and the 24 departments for which information was available were taken and the matrix of Queen-type spatial weights was also constructed with the aim of collecting interdependencies. Now, to test the null hypothesis of absence of spatial dependence, i ran several tests to discriminate between the existence of a residual spatial auto-correlation scheme (LM error and RLM error) or a spatial auto-correlation scheme in the dependent variable (LM-lag and RLM-lag).

Table 8: Spatial independence test over $\hat{A}$

	Migrant		Migrant Born	
Test	Value	P-value	Value	P-value
LM Lag	0.000	0.996	0.001	0.970
LM Error	0.009	0.924	0.136	0.713
RLM Lag	0.329	0.566	1.604	0.205
RLM Error	0.338	0.561	1.738	0.187

Table 8 shows the results of these tests, and found that the test for spatial auto-correlation in the dependent variable is not significant, therefore, there is absence of spatial dependence and no spatial auto-correlation in the dependent variable. On the other hand, the test to prove the existence of a residual spatial auto-correlation scheme (LM error), was not statistically significant at 0.05 robust and not robust.And given the results of the robust tests, i concluded that there is not possible dependence on the error given the erroneous presence of a spatially retarded endogenous variable. In other words, we confirmed that there is not spatial dependence, therefore it is not necessary to make spatial models [11].

6 *CONCLUSIONS*

This paper uses an aggregate accounting approach to analyze the relation between immigration and the productivity (and others factors productivity) of Colombian departments economies. While the aggregated nature of the data where migration is captured (GEIH) and physical capital preclude an accurate calculation of our dependent variable, several interesting results are obtained.

[11] In Appendix D you can see the Moran's I and the Spatial independence test of growth in hours worked ($\hat{x}$) and growth in intensity index $\hat{\theta}$

First, it is the confirmation of the significant effect that Venezuelan immigration has on productivity growth ($\hat{A_{dt}}$) in Colombian departments, although, unlike in developed countries, this effect is negative.This is due to the barriers presented by the Colombian labor market for Venezuelan immigrants, but without a doubt, the main reason is that such immigration is done by workers who are not sufficiently qualified and therefore there is an increase in the share of less educated workers, this confirms the mechanism of transmission of the studied effect.

Second, there is a positive and significant relationship between said immigration and growth of total hours worked ($\hat{x_{dt}}$) , this is because immigration generates a shock on the supply of labor in the labor market, there are more workers available in the market and since they are accessible labor, employers agree to hire them generating an increase in the total hours worked.

And finally, it is worth mentioning that perhaps one of the most important relationships shown in this research is that between immigration and the skill intensity index, which in turn is modified by the effect that immigration has on the shares of more and less educated workers, and therefore on the degree of skill bias in productivity. The latter relationship indicates that the more educated workers are becoming less productive relative to the less educated workers.

The results found, although different from those observed in developed countries, maintain the logic of the transmission mechanism. This confirms the hypotheses put forward. However, this research opens up many questions as to what might happen if these immigrants are placed in jobs that match their skills.

References

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134:1949–2010.

Altonji, J. and Card, D. (1989). The effects of immigration on the labor market outcomes of natives. *University of Chicago Press*, pages 201 – 234.

ANIF (2017). Productividad multifactorial y crecimiento de largo plazo en colombia. pages 1–6.

Bahar, D., Dooley, M., and Huang, C. (2018). Integración de los venezolanos al mercado laboral. *Global Economy and Development at Brookings Institute*, pages 1–23.

Barro, R. and Sala.i.Martin, X. (2003). *Economic Growth*. The MIT Press.

Borusyak, K., Hull, P., and Jaravel, X. (2018). Quasi-experimental shift-share research designs. *NATIONAL BUREAU OF ECONOMIC RESEARCH*, 134:1–43.

Calderón-Mejía, V. and Ibáñez, A. M. (2016). Labour market effects of migration-related supply shocks: Evidence from internal refugees in colombia. *Journal of Economic Geography*, 16:695–713.

Carpio, X. D. and Wagner, M. (2017). The impact of syrian refugees on the turkish economy: Regional labour market effects. *Social Sciences*, 6.

Caruso, G., Canon, C. G., and Mueller, V. (2019). Spillover effects of the venezuelan crisis: migration impacts in colombia. *Oxford Economic Papers*, pages 1–45.

D'Amuri, F. and Peri, G. (2014). Immigration, jobs, and employment protection: Evidence from europe before and during the great recession. *Journal of the European Economic Association*, 12:432–464.

Departamento Nacional de Planeación (2018). Conpes 3950. pages 1–115.

di Giovanni, J., Levchenko, A. A., and Ortega, F. (2015). A global view of cross-border migration. *Journal of the European Economic Association*, pages 168–202.

Dustmann, C., Frattini, T., and Preston, I. P. (2013). The effect of immigration along the distribution of wages. *Review of Economic Studies*, 80:145–173.

FIP (2018). Seguridad ciudadana y migración venezolana: Análisis exploratorio. Technical report, Fundación Ideas para la Paz.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2018). Bartik instruments: What, when, why, and how. No. 24408:1–50.

Hanson, G. H. (2010). *Handbook of Development Economics*, chapter Chapter 66 International Migration and the Developing World. Elsevier.

Hunt, J. and Gauthier-Loiselle, M. (2010). How much does immigration boost innovation. *American Economic Journal: Macroeconomics*, 2:31–56.

Kerr, W. R. and Lincoln, W. F. (2010). The supply side of innovation: H-1b visa reforms and u.s. ethnic invention. *Journal of Labor Economics*, 28:473–508.

Lewis, E. (2011). Immigration, skillmix, and capital skill complementarity. *Quarterly Journal of Economics*, 126:1029–1069.

Medina, C. and Posso, C. (2010). Technical change and polarization of the labor market: evidence for brazil, colombia and mexico. *Borradores de Economía; No. 614*, pages 1–52.

Migración Colombia (2019). Migración venezolana corte 30 de septiembre 2019. 5:1–8.

Morales, J. S. (2018). The impact of internal displacement on destination communities: Evidence from the colombian conflict. *Journal of Development Economics*, 131:132–150.

Orrenius, P. M. and Zavodny, M. (2012). Economic effects of migration: Receiving states. *Oxford Handbook of the Politics of International Migration.*

Peri, G. (2012). The effect of immigration on productivity: Evidence from u.s. states. *Review of Economics and Statistics*, 94:348–358.

Peri, G. (2015). Immigration and the economy of cities and regions. *Handbook of Regional and Urban Economics*, 5:625–685.

Peri, G. (2016). Immigrants, productivity, and labor markets. *Journal of Economic Perspectives*, 30:3–30.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *The Review of Economic Studies*, 87:454–486.

Universidad del Rosario (2018). Retos y oportunidades de la movilidad humana venezolana en la construcción de una política migratoria colombiana. Technical report, Universidad del Rosario.

World Bank Group (2018). Migración desde venezuela a colombia. Technical report, World Bank Group.

A Migration and Productivity: a simple model

Peri (2016) uses on a production function with constant elasticity of substitution (CES production function). The reason is that this type of production function provides a simple expression of the marginal (logarithmic) productivity of each skill as a function of the supply of the same skill, of simple aggregations of other skill supplies and of a small number of parameters; in short, it presents a simplicity in interpreting the relationships and the mechanism of impact of immigration on productivity and other variables included in the production function. In this way, it is useful to describe some details of how the nested ESC production function approach can be used to estimate elasticity parameters of the effects of the increase in labor supply through immigrants on productivity; the most relevant characteristics used to organize cells in the nested-CES framework have been education levels, age groups (or experience groups), and nativity groups (foreign–native). These have provided the grid to organize workers into cells. Adopting a CES structure, one could represent a production function with a small number of parameters. And one would be able to estimate those parameters using the whole country as relevant area, simply exploiting the variation of immigrant supply over time and across skill cells (Peri, 2015). Below is a step-by-step sketch of the mathematical procedure of the model:

$$Y_{st} = K_{st}^{\alpha}[X_{st}A_{st}\theta(h_{st})]^{1-\alpha}$$

Here, we divide the above equation by N_{st}

$$\begin{aligned}
\frac{Y_{st}}{N_{st}} &= \frac{K_{st}^{\alpha}[X_{st}A_{st}\theta(h_{st})]^{1-\alpha}}{N_{st}} \\
y_{st} &= \frac{K_{st}^{\alpha}}{N_{st}}\frac{[X_{st}A_{st}\theta(h_{st})]}{[X_{st}A_{st}\theta(h_{st})]^{\alpha}} \\
y_{st} &= \frac{K_{st}^{\alpha}}{[X_{st}A_{st}\theta(h_{st})]^{\alpha}}\frac{[X_{st}A_{st}\theta(h_{st})]}{N_{st}} \\
y_{st} &= \left(\frac{K_{st}}{K_{st}^{\alpha}[X_{st}A_{st}\theta(h_{st})]^{1-\alpha}}\right)^{\frac{\alpha}{1-\alpha}}\frac{[X_{st}A_{st}\theta(h_{st})]}{N_{st}} \\
y_{st} &= \left(\frac{K_{st}}{Y_{st}}\right)^{\frac{\alpha}{1-\alpha}}[x_{st}A_{st}\theta(h_{st})]
\end{aligned}$$

And in order to obtain the growth rate of the product, we take logarithms on both sides of the above equation, thus obtaining the following equation:

$$\hat{Y_{st}} = \hat{N_{st}} + \hat{y_{st}} = \hat{N_{st}} + \frac{\alpha}{1-\alpha}[\frac{\hat{K_{st}}}{Y_{st}}] + \hat{A_{st}} + \hat{x_{st}} + \hat{\theta_{st}}$$

However, as mentioned in the research, there are two fundamental variables that are not observable, both productivity and skill bias in productivity. The goal of the following procedure is to make both unobservable variables dependent on variables and coefficients that are observable in our database and then to calculate them. In this way, (Peri, 2012) adjusts the ratio of wages per hour worked of the more and less educated workers, and equals it to the ratio of their marginal productivities, obtaining the following equation:

$$\frac{W_{st}^{H}}{W_{st}^{L}} = \left(\frac{\beta_{st}}{1-\beta st}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{h_{st}}{1-h_{st}}\right)^{-\frac{1}{\sigma}}$$

Then, we proceed to clear β_{st}

$$\begin{aligned}
\left(\frac{\beta_{st}}{1-\beta st}\right)^{\frac{\sigma-1}{\sigma}} &= \frac{W_{st}^{H}\,(h_{st})^{\frac{1}{\sigma}}}{W_{st}^{L}\,(1-h_{st})^{\frac{1}{\sigma}}} \\
\beta_{st}^{\frac{\sigma-1}{\sigma}} &= \frac{W_{st}^{H}\,(h_{st})^{\frac{1}{\sigma}}}{W_{st}^{L}\,(1-h_{st})^{\frac{1}{\sigma}}}(1-\beta st)^{\frac{\sigma-1}{\sigma}} \\
\beta_{st}^{\frac{\sigma-1}{\sigma}}\left(1+\frac{W_{st}^{H}\,(h_{st})^{\frac{1}{\sigma}}}{W_{st}^{L}\,(1-h_{st})^{\frac{1}{\sigma}}}\right) &= \frac{W_{st}^{H}\,(h_{st})^{\frac{1}{\sigma}}}{W_{st}^{L}\,(1-h_{st})^{\frac{1}{\sigma}}} \\
\beta_{st}^{\frac{\sigma-1}{\sigma}}\left(\frac{W_{st}^{L}\,(1-h_{st})^{\frac{1}{\sigma}}+W_{st}^{H}\,(h_{st})^{\frac{1}{\sigma}}}{W_{st}^{L}\,(1-h_{st})^{\frac{1}{\sigma}}}\right) &= \\
\frac{W_{st}^{H}\,(h_{st})^{\frac{1}{\sigma}}}{W_{st}^{L}\,(1-h_{st})^{\frac{1}{\sigma}}} \\
\beta_{st}^{\frac{\sigma-1}{\sigma}} &= \frac{W_{st}^{H}\,(h_{st})^{\frac{1}{\sigma}}}{W_{st}^{L}\,(1-h_{st})^{\frac{1}{\sigma}}+W_{st}^{H}\,(h_{st})^{\frac{1}{\sigma}}}
\end{aligned}$$

Thus obtaining the equation of the skill bias in productivity:

$$\beta_{st} = \frac{(W_{st}^{H})^{\frac{\sigma}{\sigma-1}}(hst)^{\frac{1}{\sigma-1}}}{(W_{st}^{L})^{\frac{\sigma}{\sigma-1}}(1-h_{st})^{\frac{1}{\sigma-1}}+(W_{st}^{H})^{\frac{\sigma}{\sigma-1}}(h_{st})^{\frac{1}{\sigma-1}}}$$

However, we need the skill intensity index to be dependent on the skill bias, so we replace the previous equation in equation 2, thus obtaining the following procedure:

$$\theta(h_{st}) = \left(\frac{W_{st}^{H}\left(h_{st}\right)}{W_{st}^{L}\left(1-h_{st}\right)^{\frac{1}{\sigma}}+W_{st}^{H}\left(h_{st}\right)^{\frac{1}{\sigma}}}+\right.$$

$$\left(\left(\frac{(W_{st}^{L})^{\frac{\sigma}{\sigma-1}}(1-h_{st})^{\frac{1}{\sigma-1}}}{(W_{st}^{L})^{\frac{\sigma}{\sigma-1}}(1-h_{st})^{\frac{1}{\sigma-1}}+(W_{st}^{H})^{\frac{\sigma}{\sigma-1}}(h_{st})^{\frac{1}{\sigma-1}}}\right)X\right.$$

$$\left.(1-h_{st}))^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$\theta(h_{st}) = \left(\frac{W_{st}^{H}\left(h_{st}\right)}{W_{st}^{L}\left(1-h_{st}\right)^{\frac{1}{\sigma}}+W_{st}^{H}\left(h_{st}\right)^{\frac{1}{\sigma}}}+\right.$$

$$\left.\frac{W_{st}^{L}\left(1-h_{st}\right)}{W_{st}^{L}\left(1-h_{st}\right)^{\frac{1}{\sigma}}+W_{st}^{H}\left(h_{st}\right)^{\frac{1}{\sigma}}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$\theta(h_{st}) = \left(\frac{W_{st}^{H}\left(h_{st}\right)+W_{st}^{L}\left(1-h_{st}\right)}{W_{st}^{L}\left(1-h_{st}\right)^{\frac{1}{\sigma}}+W_{st}^{H}\left(h_{st}\right)^{\frac{1}{\sigma}}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$\theta(h_{st}) = \frac{((W_{st}^{H})(hst)+(W_{st}^{L})(1-hst))^{\frac{\sigma}{\sigma-1}}}{(W_{st}^{L})^{\frac{\sigma}{\sigma-1}}(1-h_{st})^{\frac{1}{\sigma-1}}+(W_{st}^{H})^{\frac{\sigma}{\sigma-1}}(h_{st})^{\frac{1}{\sigma-1}}}$$

Now that all the variables are observable, we take equation 1 and clear for productivity:

$$Y_{st} = K_{st}^{\alpha}[X_{st}A_{st}\theta(h_{st})]^{1-\alpha}$$

$$\left(\frac{Y_{st}}{K_{st}^{\alpha}}\right)^{\frac{1}{1-\alpha}} = X_{st}A_{st}\theta(h_{st})$$

$$A_{st} = \left(\frac{Y_{st}^{\frac{1}{1-\alpha}}K_{st}^{-\frac{\alpha}{1-\alpha}}}{X_{st}}\right)*\frac{1}{\theta(h_{st})}$$

Getting the following equation once we replace $\theta(h_{st})$:

$$A_{st} = \left(\frac{Y_{st}^{\frac{1}{1-\alpha}}K_{st}^{-\frac{\alpha}{1-\alpha}}}{X_{st}}\right)\times$$

$$\frac{(W_{st}^{L})^{\frac{\sigma}{\sigma-1}}(1-h_{st})^{\frac{1}{\sigma-1}}+(W_{st}^{H})^{\frac{\sigma}{\sigma-1}}(h_{st})^{\frac{1}{\sigma-1}}}{((W_{st}^{H})(hst)+(W_{st}^{L})(1-hst))^{\frac{\sigma}{\sigma-1}}}$$

B Data summary

Table 9 shows each of the variables taken into account in the research with their respective description and source

Table 9: Description variables

Variable	Description	Source
$\hat{y}$	Output growth per worker	DANE
$\hat{k}$	Capital growth per worker	DANE
$\hat{A}$	Productivity growth	Own calculation with DANE and GEIH data
$\hat{h}$	Growth in the share of total hours worked supplied by the high educated workers	GEIH
$\widehat{(1-h)}$	Growth in the share of total hours worked supplied by the less educated workers	GEIH
H	Total hours worked by Natives and Venezuelan high educated workers(total hours in a year)	GEIH
L	Total hours worked by Natives and Venezuelan less educated workers(total hours in a year)	GEIH
$\hat{x}$	Growth of total hours worked by workers(total hours in a year)	GEIH
X	Total hours worked	GEIH
X^L	Total hours worked by less educated workers	GEIH
X^H	Total hours worked by high educated workers	GEIH
W_M^H	Average hourly wage by high educated wokers (with Migrant variables)	GEIH
W_M^L	Average hourly wage by Less educated wokers (with Migrant variables)	GEIH
W_{MB}^H	Average hourly wage by high educated wokers (with Migrant Born variables)	GEIH
W_{MB}^L	Average hourly wage by Less educated wokers (with Migrant Born variables)	GEIH
N	Total number of people employed	GEIH
N^c	Total natives employed	GEIH
N^m	Total migrants employed (Migrant variable)	GEIH
N^{mb}	Total migrants employed (Migrant Born variable)	GEIH
$\hat{\theta}$	Gowth of skill intensity index	Own calculation with GEIH data
$\hat{\beta}$	Growth of skill bias of productivity	Own calculation with GEIH data
	Parámetros	
α	Share of capital stock	Medina and Posso (2010)
σ	Sustitution elasticity between less and high educated workers	Medina and Posso (2010)

C First stage Regression

The following tables correspond to the first-stage regression of the 2SLS models.

Table 10: First stage Regression (Migrant)

	(1)
Explanatory variable	Instrumental Variable
$\frac{\Delta N^f_{st}}{N_{st}}$	5.27e-08***
	(9.39e-09)
R-sq	
Within	0.1414
Between	0.7013
Overall	0.2124
Observations	120
F-Test	31.52

Note: The explanatory variable is immigration as a percentage of initial employment. Each cell is the result of a separate regression. The units of Observations are Colombian departments in year 2014–2018. Each regression includes department fixed effects. The method of estimation is least squares with observations weighted by the employment of the state. Errors in parentheses are heteroskedasticity robust and clustered by departments. The calculated variables use the assumption that $\sigma = 1.47$ and $\alpha = 0.3$. *Significant 5%, **1%, ***0.1%.

Table 11: First stage Regression (Migrant Born)

	(1)
Explanatory variable	Instrumental Variable
$\frac{\Delta N^f_{st}}{N_{st}}$	9.48e-08***
	(1.11e-08)
R-sq	
Within	0.6058
Between	0.7710
Overall	0.5447
Observations	120
F-Test	73.15

Note: The explanatory variable is immigration as a percentage of initial employment. Each cell is the result of a separate regression. The units of Observations are Colombian departments in year 2014–2018. Each regression includes department fixed effects. The method of estimation is least squares with observations weighted by the employment of the state. Errors in parentheses are heteroskedasticity robust and clustered by departments. The calculated variables use the assumption that $\sigma = 1.47$ and $\alpha = 0.3$. *Significant 5%, **1%, ***0.1%.

D Moran's I and independence tests

Table 12: Moran's I over $\hat{x}$

	Migrant		Migrant Born	
Year	Moran's I	P-value	Moran's I	P-value
2014	1.370	0.171	1.089	0.276
2015	1.051	0.293	0.542	0.588
2016	1.279	0.201	0.917	0.359
2017	1.511	0.131	1.373	0.170
2018	1.616	0.106	1.528	0.126

Table 13: Spatial independence test over $\hat{x}$

	Migrant		Migrant Born	
Test	Value	P-value	Value	P-value
LM Lag	1.176	0.278	1.311	0.252
LM Error	1.132	0.287	1.296	0.255
RLM Lag	0.057	0.811	0.019	0.890
RLM Error	0.014	0.907	0.004	0.947

Table 14: Moran's I over $\hat{\theta}$

	Migrant		Migrant Born	
Year	Moran's I	P-value	Moran's I	P-value
2014	-0.286	1.225	-0.058	1.047
2015	-0.102	1.082	0.296	0.767
2016	-0.500	1.383	-0.608	1.457
2017	-0.957	1.662	-0.925	1.645
2018	-0.493	1.378	-0.534	1.407

Table 15: Spatial independence test over $\hat{\theta}$

	Migrant		Migrant Born	
Test	Value	P-value	Value	P-value
LM Lag	8.477	0.004	7.690	0.006
LM Error	11.804	0.001	10.761	0.001
RLM Lag	2.995	0.084	4.632	0.031
RLM Error	6.322	0.012	7.704	0.006