

THREE ESSAYS ON IMMIGRATION AND AGRICULTURAL
ECONOMICS

A Dissertation

Presented to the Faculty of the Graduate School
of Cornell University

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by

Andrés Mauricio Castaño Zuluaga

August 2023

© 2023 Andrés Mauricio Castano Zuluaga

THREE ESSAYS ON IMMIGRATION AND AGRICULTURAL ECONOMICS

Andrés Mauricio Castaño Zuluaga, Ph. D.

Cornell University 2023

Even though 35% of the migration is between developing countries, most of the literature on the impact of international migration on natives' fiscal contributions and internal mobility decisions is concentrated among OECD countries. In this dissertation, in the first two chapters, I investigate the effects of the arrival of Venezuelan immigrants to Colombia on the fiscal contributions and internal mobility of natives. The third chapter explores the potential for chocolate made from CCN-51 to substitute high-quality chocolate.

In the first chapter, along with Carlos Mesa Guerra, we reexamine the effect of immigration on public finances by accounting for second-order effects. We exploit exogenous variation in immigration across Colombian metropolitan areas between 2013 and 2018, resulting from the significant increase in Venezuelan immigrants and instrument immigrants' residential locations using pre-existing settlement patterns and the distance between origin-destination flows. Our findings indicate that immigration did not reduce natives' average fiscal contributions.

In the second chapter, I exploit plausible exogenous variation in immigration from Venezuela across Colombian states to study the effect of changes in immigration inflows on changes in natives' net migration rates and their components (in and out-migration rates) between 2014-2019. I found no evidence that immigration from Venezuela induces changes in net-migration rates across Colombian states. I also found evidence that although changes in immigration rates are positively associated with natives' out-migration and negatively associated with natives' in-migration rates, the effects are not distinguishable from zero once confidence sets robust to weak instruments are constructed.

In the third chapter, I used an auction experiment to study the hedonic properties and the factors affecting the WTP for Ecuadorian chocolate produced with different cacao, including the CCN-51 cacao hybrid. We find that smell ratings favored the highest quality cacao compared to the CCN-51 hybrid; however, there were no differences in taste ratings. We also found no difference in mean WTP between the national variety and the CCN-51 cacao hybrid, even when information on the quality and nutritional value of the cacao was provided. Results suggest that the expansion in the use of CCN-51 cacao in chocolate production is a plausible alternative to deal with the potential shortage of chocolate.

BIOGRAPHICAL SKETCH

Andrés Castaño was born in 1988 in San Rafael, a municipality in Antioquia, Colombia. He attended Kindergarten and primary education in a public primary school in Cartagena, Colombia. He then completed high school at Colegio Militar Almirante Colon at 17. He obtained his B.A degree in Economics in 2012 at Universidad de Cartagena, Cartagena, Colombia. Before completing his degree, he worked as a Research Assistant for the Centre for Regional Economics Research at the Central Bank of Colombia, Cartagena, Colombia. After graduating from college, Mr. Castaño moved to Antofagasta, Chile, to complete his Master of Science in Regional Science under the supervision of Professor Dusan Paredes. During his studies, he worked as a Research Assistant for the Institute of Applied Regional Economics at Universidad Católica del Norte. After obtaining his master's degree, he worked as an instructor for core economic courses at the undergraduate level at the same university. Before beginning the Ph.D. program at Cornell University, he was an Economic Consultant for the office of the Vice Provost at the Universidad Católica del Norte. In this position, Mr. Castaño carried out econometric analyses and data management to identify students at risk of dropping out and to design policies to reduce students' desertion. Through this and previous experiences, Mr. Castaño developed an interest in Regional Economics and

Applied Econometrics, motivating him to pursue the Ph.D. program in Regional Science at Cornell University. While completing his Ph.D. required courses, he developed an interest in understanding the socioeconomic effects of Venezuelan immigration in Colombia, which led to research ideas later explored as part of his Ph.D. dissertation. Mr. Castaño is now an Assistant Teaching Professor in Business Analytics at Carnegie Mellon University in Doha, Qatar. Mr. Castaño enjoys playing soccer, socializing with friends, and reading non-fiction in his free time.

To my parents, family, and friends

ACKNOWLEDGMENTS

I am grateful to my dissertation committee for their patience and insights. I thank Professor Kieran Donaghy, the chairman of my committee, for his supervision throughout the writing of my dissertation. To Professor Miguel Gomez, a committee member, I am infinitely grateful for his support and encouragement during difficult times and for all the help getting the data and writing chapter three. I owe many thanks to Professor Lawrence Kahn for helping me with outstanding feedback on the chapters related to immigration. Thank you for making my defense enjoyable and for your brilliant comments and suggestions.

In addition to my committee members, I have benefited from many other mentors I had previously, namely Marcelo Lufin, Miguel Atienza, and David Silbey. I thank them all for their continuous support during my Ph.D. studies, related research, motivation, and immense knowledge.

Also, I would like to thank my family for all their love and encouragement. To my parents for their continuous encouragement and love. My sisters for always been there for me. I wish to thank my friends in Colombia, particularly my friend from life, Leonardo Castillo, who's always been there for me. To my friends during my stay at Cornell, especially all the group from The Colombian Association of Students at Cornell, for making my time in Ithaca a wholesome experience. In particular Ana

Porras, Javier Jaimes, Jose Bastidas, Rodrigo Molano, and Juan Jose Silva. I also want to thank all the incredible human beings at Cornell in the Washington Program, David Silbey, Carol Fields, Dan Siegel, Brittany Paylor, Elizabeth Burgess, and John Cawley, for their encouraging words and for allowing me to work for the CIW program. It's been one of the most impactful experiences in my career, and it allowed me to rediscover my passion for working with students and provided all the resources to complete my dissertation. Thank you!

Finally, I would like to thank my sponsor Becas Chile, from Ministerio de Educación, Government of Chile, because, without their financial support, this Ph.D. would have never happened.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH.....	V
ACKNOWLEDGMENTS.....	VIII
TABLE OF CONTENTS	X
LIST OF FIGURES	XII
LIST OF TABLES.....	XIII
CHAPTER 1: THE FISCAL IMPACT OF IMMIGRATION: LABOR DISPLACEMENT, WAGES, AND THE ALLOCATION OF PUBLIC SPENDING	1
1.1 INTRODUCTION	1
1.2 THEORETICAL FRAMEWORK.....	10
1.3 DATA AND BACKGROUND	14
1.3.1 Data	14
1.3.2 Background.....	18
1.4 EMPIRICAL FRAMEWORK.....	21
1.4.1 Estimation Strategy	21
1.4.2 Identification.....	24
1.5 RESULTS.....	31
1.5.1 Effect on Natives' Fiscal Contributions	31
1.5.2 Robustness Checks.....	34
1.5.3 Differences by Revenue and Expenditure Group.....	40
1.6 MECHANISMS	42
1.6.1 Labor Market Effects	42
1.6.2 Property Values and Rents.....	55
1.6.3 Cost Effect.....	60
1.6.4 Composition and Service Use Effect.....	62
1.7 DISCUSSION	67
1.8 CONCLUDING REMARKS.....	70
REFERENCES	73
APPENDIX.....	78
A. Additional Figures and Tables	78
B. Data Appendix	88
CHAPTER 2: IMMIGRATION AND NATIVES' INTERNAL MOBILITY IN THE DEVELOPING WORLD: EVIDENCE FROM THE VENEZUELAN IMMIGRATION TO COLOMBIA BETWEEN 2014-2019	101
2.1 INTRODUCTION	101
2.2 THEORETICAL FRAMEWORK.....	110
2.3 DATA AND DESCRIPTIVE STATISTICS.....	111
2.4 EMPIRICAL STRATEGY.....	119
2.5 RESULTS	130
2.5.1 Effects on population changes and migration rates	130
2.5.2 Robustness Checks.....	140
2.6 CONCLUDING REMARKS.....	143
REFERENCES	146

APPENDIX.....	150
CHAPTER 3: IS THE CCN-51 CACAO HYBRID THE SOLUTION FOR THE POTENTIAL SHORTAGE OF CHOCOLATE? A HEDONIC AND WILLINGNESS TO PAY STUDY FOR DIFFERENT VARIETIES OF ECUADORIAN CHOCOLATE	152
3.1 INTRODUCTION	152
3.2 BACKGROUND ON THE COCOA AND CHOCOLATE MARKETS	159
3.3 EXPERIMENTAL DESIGN	163
3.4 DATA AND EMPIRICAL SPECIFICATION	167
3.5 RESULTS AND DISCUSSION	178
3.6 CONCLUSION	183
REFERENCES	185
APPENDIX.....	190
A. Slides explaining the experiment to the NO INFO group	190
B. Slides explaining the experiment to the INFO group	196
C. Consent Form	203
D. Instructions Auction Experiment	205
E. First Trial Auction: Bidding for a One Dollar bill	205
F. Second Trial Auction: Bidding for a Candy Bar	206
G. Instruction-Chocolate Auctions	207
H. Exit Survey	209

LIST OF FIGURES

FIGURE 1.1 MIGRANT STOCK, APRIL 2013 - MARCH 2019-----	21
FIGURE 1.2 PRE-TRENDS USING THE EXPOSURE SHARES IN 2005-----	29
FIGURE 1.3 CHANGE IN NATIVES' FISCAL CONTRIBUTIONS BY REVENUE AND EXPENDITURE GROUP (<i>COP\$ MILLION, 2018 EQUIVALENT</i>)-----	41
FIGURE 1.4 EFFECT OF IMMIGRATION ON MAIN LABOR MARKET OUTCOMES-----	45
FIGURE 1.5 EFFECT OF VENEZUELAN IMMIGRATION ALONG NATIVES' WAGE DISTRIBUTION --	52
FIGURE 1.6 EFFECT OF IMMIGRATION ON THE INTENSIVE MARGIN AND THE QUALITY OF JOBS	55
FIGURE 1.7 COMPOSITION AND SERVICE USE EFFECT-----	65
FIGURE 1.8 FISCAL CONTRIBUTIONS, IMMIGRANT SHOCK, AND ENERGY CYCLE-----	78
FIGURE 1.9 DIRECTED ACYCLIC GRAPH (DAG) OF THE EFFECT OF IMMIGRATION ON NATIVES' FISCAL CONTRIBUTIONS-----	78
FIGURE 1.10 YEAR-BY-YEAR ESTIMATES OF THE EFFECT OF IMMIGRATION ON NATIVES' FISCAL CONTRIBUTIONS-----	79
FIGURE 1.11 ESTIMATES OF THE EFFECT OF IMMIGRATION ON NATIVES' FISCAL CONTRIBUTIONS RELATIVE TO 2013-----	80
FIGURE 1.12 EFFECT OF IMMIGRATION ON MAIN LABOR MARKET OUTCOMES BY SEX-----	81
FIGURE 1.13 EFFECT OF IMMIGRATION ON MAIN LABOR MARKET OUTCOMES BY SKILL-----	82
FIGURE 1.14 EFFECT OF IMMIGRATION ON WAGES BY SECTOR-----	83
FIGURE 1.15 EFFECT OF IMMIGRATION ON WAGES BY TYPE OF JOB-----	84
FIGURE 1.16 ACTUAL VS. PREDICTED POSITION OF VENEZUELAN IMMIGRANTS IN NATIVES' WAGE DISTRIBUTION-----	85
FIGURE 2.1 CHANGES IN NATIVE POPULATION VS. CHANGES IN VENEZUELA IMMIGRATION, 2017 AND 2019-----	117
FIGURE 2.2 CHANGES IN NATIVES' IN AND OUT MIGRATIONS RATES VS. CHANGES IN IMMIGRATION FROM VENEZUELA, 2017 AND 2019-----	118
FIGURE 2.3 PRE-TRENDS FOR THE EXPOSURE CHARES-----	130
FIGURE 2.4 MONTHLY STOCK OF IMMIGRANTS FROM VENEZUELA ACROSS COLOMBIAN MSAs -----	150
FIGURE 2.5 EVOLUTION IN THE SHARE OF IMMIGRATION FROM VENEZUELA IN THE NATIVE WORKING-AGE POPULATION 2013-3 – 2019-12-----	151
FIGURE 3.1 COMPARISON OF WTP AMONG TREATMENT GROUPS-----	169
FIGURE 3.2 HEDONIC PROPERTIES OF THE DIFFERENT VARIETIES OF CHOCOLATE BY TREATMENT GROUP-----	172
FIGURE 3.3 WILLINGNESS TO PAY ESTIMATES USING RANDOM-EFFECTS TOBIT AND OLS MODELS-----	180

LIST OF TABLES

TABLE 1.1 DESCRIPTIVE STATISTICS FOR 2013 AND 2018-----	18
TABLE 1.2 RELATIONSHIP BETWEEN EXPOSURE SHARES AND LOCAL CHARACTERISTICS-----	27
TABLE 1.3 EFFECT OF IMMIGRATION ON NET FISCAL CONTRIBUTIONS – FIRST DIFFERENCES--	32
TABLE 1.4 EFFECT OF IMMIGRATION ON REVENUES AND EXPENDITURES-----	35
TABLE 1.5 ROBUSTNESS ESTIMATES OF THE EFFECT OF IMMIGRATION ON NATIVES’ FISCAL CONTRIBUTIONS – RESTRICTED BUDGET-----	39
TABLE 1.6 EFFECT OF IMMIGRATION ON PROPERTY VALUES AND RENTS-----	59
TABLE 1.7 ROBUSTNESS ESTIMATES OF THE EFFECT OF IMMIGRATION ON NATIVES’ FISCAL CONTRIBUTIONS – ALL BUDGET-----	86
TABLE 1.8 EFFECT OF IMMIGRATION ON NATIVES’ OUTFLOWS-----	86
TABLE 1.9 LIST OF GOVERNMENT REVENUES-----	94
TABLE 1.10 LIST OF GOVERNMENT EXPENDITURES BY SUB-FUNCTION (UN COFOG-DANE) -----	98
TABLE 2.1 DESCRIPTIVE STATISTICS-----	114
TABLE 2.2 RELATIONSHIP BETWEEN EXPOSURE SHARES AND LOCAL CHARACTERISTICS IN 2005 AND 2013-----	128
TABLE 2.3 EFFECT OF IMMIGRATION FROM VENEZUELA ON NATIVE POPULATION CHANGES ACROSS STATES, 2013-2019-----	132
TABLE 2.4 EFFECT OF IMMIGRATION FROM VENEZUELA ON NATIVES’ IN AND OUT-MIGRATION RATES ACROSS STATES, 2013-2019-----	135
TABLE 2.5 ROBUSTNESS EXERCISES FOR THE EFFECT OF CHANGES IN IMMIGRATION FROM VENEZUELA ON CHANGES IN NATIVES’ NET MIGRATION RATES AND ITS COMPONENTS---	141
TABLE 3.1 EXPERIMENTAL DESIGN: INFORMATION PROVIDED FOR THE THREE VARIETIES OF ECUADORIAN CHOCOLATE-----	166
TABLE 3.2 DESCRIPTIVE STATISTICS FOR WILLINGNESS TO PAY (WTP), DEMOGRAPHICS, AND HABIT PURCHASE DATA-----	168
TABLE 3.3 BALANCE IN THE EXPLANATORY VARIABLES. MEAN DIFFERENCES BY INFORMATION TREATMENT ACROSS EXPLANATORY VARIABLES-----	170
TABLE 3.4 DIFFERENCES IN HEDONIC RATINGS AND WTP AMONG CHOCOLATE VARIETIES BY INFORMATION TREATMENT-----	176
TABLE 3.5 AVERAGE ADJUSTED PREDICTIONS (AAPs) BY CHOCOLATE TYPE AND TREATMENT -----	182

CHAPTER 1

THE FISCAL IMPACT OF IMMIGRATION: LABOR DISPLACEMENT, WAGES, AND THE ALLOCATION OF PUBLIC SPENDING

1.1 Introduction

With rising debt and deficits across many developed and developing countries, immigrants are often perceived to be a burden for public finances among native-born residents of immigrant-receiving countries. This has been the case even when the empirical evidence has shown for several countries that the net (direct) fiscal impact of immigration is typically less than 0.5% of GDP for a given year, regardless of whether it is positive or negative (OECD, 2013; Dustmann & Frattini, 2014; National Academies of Sciences, Engineering, and Medicine, 2017). For instance, total immigration in Colombia is estimated to generate a small fiscal benefit (Mesa-Guerra & Ramírez-Tobón, 2022), but consistently more than half of residents consider immigration to be a fiscal burden. Does the perception of natives reflect other indirect effects of immigration that affect their tax contributions and welfare benefits? Are these effects being omitted in the conventional fiscal accounting exercises, overestimating the positive fiscal impact of immigrants—or underestimating their effect when they don't do so well?

The conventional accounting approach to estimate the impact of immigration on public finances usually ignores the effects that immigration has on natives' labor market outcomes or property values, or in the allocation of government expenditures, to name a few examples. Results derived from first-order fiscal effects, *i.e.*, the difference between the contributions immigrants make in the form of taxes and the government expenditures they receive, assume no impact on the wages and employment of natives,¹ changes in the return to capital, fiscal benefits from economic growth, or lower per capita costs of providing public goods that are distributed over a larger population. By ignoring these second-order effects, it assumed that the economy is able to absorb immigration without affecting the fiscal contributions of natives.

In this paper, we reexamine the effect of immigration on public finances by addressing these second-order effects. Specifically, we look at labor displacement, changes in factor prices (labor and capital), and

¹ The impact of immigration on natives' wages and employment has been widely studied in the empirical literature in labor economics. Studies that exploit variation in immigrant inflows across education-experience cells (Borjas, 2003) tend to find negative—and larger—effects than those studies that exploit variation across geographical units (Altonji & Card, 1991; Dustmann *et al.*, 2013; Foged & Peri, 2016), or that use a mixture of the skill-cell and spatial approach (Card, 2001, 2009). Our approach in this paper is closer in spirit to the latter. Another strand of literature using a structural approach has found positive effects for native workers (Ottaviano & Peri, 2012). More recently, Dustmann *et al.* (2017) using a natural experiment from a change in a commuting policy along the German-Czech border show a moderate decline in wages but a sharp decline in employment at the local level for natives. Dustmann *et al.* (2016) shows that the different empirical approaches identify conceptually different parameters. While the skill-cell and mixed approaches identify *relative* effects of immigration on natives by experience within each education group, the spatial approach recovers the *total* effect on the native group of interest.

changes in the allocation and cost of providing public services. To guide our empirical analysis, we present a simple theoretical framework that allows us to derive the implications of changes in population driven by increasing immigration on natives' net fiscal contributions. We use this framework to explain differences in fiscal contributions for natives and set the ground for exploring the mechanisms in place. We show that the effect of increasing immigrant inflows on the net fiscal contributions of native workers depends on three terms. First, a *factor price effect* that captures changes in taxes paid that result from both labor displacement and variations in wages, as part of the labor market adjustment, and changes in revenue from capital income. Second, the change in the cost of government service provision. This *cost effect* may be driven by the mix of goods and services provided (pure, congested, or private goods), changes in the population composition, and the overall increase in the local population. Third, changes in the *composition* of government expenditures and differences in the demand for government services between natives and immigrants.

Empirically, we exploit exogenous variation in immigration across Colombian cities with their metropolitan areas between 2013 and 2018, resulting from the significant increase in Venezuelan immigrants. This includes both native-born returnees and Venezuelan-born immigrants. Based on individual-level estimates of tax payments and benefits received from the use of public goods and services—such as education or health

care—and welfare, we present reduced-form evidence of the effect of changes in the stock of immigrants on natives’ per capita net fiscal contributions. We then explore the three mechanisms that are consistent with our theoretical framework.

To address the potential endogeneity from immigrants’ self-selecting into local labor markets and measurement error from small sample sizes used to construct the fraction of immigrants at the regional level, we follow the literature and use two shift-share instruments. The first instrument uses predicted immigrant inflows based on pre-existing settlement patterns from the 2005 Census (Altonji & Card, 1991; Card, 2001). This instrument is constructed using historic settlement shares, using the well-known fact that immigrants tend to locate disproportionately near previous immigrants with similar backgrounds. The second instrument uses the distance between origin-destination flows. The instrument predicts the number of immigrants that would have arrived in a particular city and year based on the regional origin-distribution of aggregate inflows and the travel distance between the centroid of the capital for each Venezuelan province and the centroid of the capital city in each Colombian department (Del Carpio & Wagner, 2015; Caruso *et al.* 2019; Delgado-Prieto, 2022). Since Colombia and Venezuela share a 2,200 km border with multiple regular and irregular border crossings, distance is a key determinant of immigrants’ location decisions.

Our results show no evidence that increasing immigrant inflows to Colombia lowered the average per capita contributions of natives between 2013 and 2018. These effects stay unchanged when we exclude revenues and expenditures not derived from individual behavior, such as corporate taxes, gross operating surplus, royalties, ‘pure’ public goods, and debt service. However, when we run yearly regressions separately, we see a small negative effect in 2017, when the pace of immigrant inflows increased exponentially. Nevertheless, the implied effects are relatively small, roughly equivalent to about 0.3% of GDP per capita.

We also find no indication of a significant effect on average tax contributions and a positive effect on expenditures only when we allow for differential trends in fiscal contributions across areas. The latter, however, is not distinguishable from zero when we construct identification-robust Anderson–Rubin confidence intervals. Our estimates are robust to controlling for endogenous changes in the local population, using different instruments, using the pooled sample, or allowing for dynamic effects.

When we disaggregate the effects by types of revenues and expenditures, we find a slight drop in indirect tax contributions and an increase in health care benefits. The former is explained by a decline in wages for low-skilled workers, mainly at the bottom of the wage distribution. However, we believe these effects are too small to drive large

changes in fiscal contributions, mainly since we do not find evidence of labor displacement. In addition, we find no evidence of costs being distributed among a larger population and no significant change in local spending that could be explained by increasing immigrant inflows. However, we find evidence that changes in the fraction of immigrants are associated with an increase in natives' welfare take-up of 1.7 percentage points, driven by a higher probability of receiving subsidized health.

These results have four implications. First, the composition of public expenditures is biased towards services that tend to increase with population. Second, local governments have had to stretch their resources to meet the increasing demand as they face low budget flexibility. Third, the increase in demand for welfare benefits has put more pressure on resources coming from the National government. Fourth, general equilibrium effects are not large enough to influence changes in fiscal contributions. Our findings accrue to small aggregate effects and, therefore, net fiscal estimates presented here are not a result of indirect effects canceling out.

This paper builds on the extensive literature on the fiscal effects of immigration. While the evidence on the direction of the effect is mixed, as one would expect from the analysis of very different settings, the evidence on the size of the effect has shown it to be fairly small compared to the size

of the overall fiscal imbalance and the economy (OECD, 2013). Our paper is more closely related to the strand of research that has used a static fiscal accounting framework, *i.e.*, estimating the differences between the value of taxes and other contributions immigrants make to revenues and the value of expenditures received in the form of public services and welfare. Dustmann & Frattini (2014), using a repeated cross-sectional approach, show that immigrants to the UK since 2000 have contributed more to taxes than they have added to expenditures. In an extensive study of immigration in the U.S., the report by the National Academies of Sciences, Engineering, and Medicine (2017) shows that while first-generation immigrants (*i.e.*, foreign-born immigrants) have had lower fiscal contributions relative to the other generations, mainly at the state and local levels, second-generation immigrants (*i.e.*, immigrants' children) are among the larger fiscal contributors in the population. Recent evidence for developing countries finds considerable variation in the relative net fiscal contributions between immigrants and natives and shows that the per-capita effect is relatively large (OECD/ILO, 2018).

Using a similar approach, Mesa-Guerra & Ramírez-Tobón (2021) provide evidence of the fiscal effect of the recent Venezuelan immigration to Colombia. Their results indicate that the lower fiscal contributions of immigrants—relative to natives—are driven by recent arrivals who have less attachment to the labor market. However, immigrants that have been

in the country for more than a year show equal or slightly higher per capita fiscal contributions than natives after controlling for differences in the age profile and other individual characteristics.

Our contribution to this literature is to account for second-order effects. We account for changes in natives' employment, wages, and in the cost and allocation of public spending at the local level. While immigration can improve natives' fiscal position in the sense that 'pure' public goods can be provided at the same level while sharing the costs among the whole population—at least in the short run when the marginal cost is lower than the average cost, it can have a negative effect by reducing tax payments from natives as a result of job loss or lower wages. Our findings show that the wage response across skill groups is not uniform. We find negative and significant effects on wages between the 15th and 50th percentiles and for low-skilled and self-employed workers. However, because the tax contributions of these groups tend to be low, it is not unusual to see no change in the overall fiscal effect as a response to immigration.

Finally, a different literature has documented the relationship between states' or cities' ethnic or racial composition and public spending. For example, Alesina *et al.* (1999) show that the allocation of spending to productive public goods (*e.g.*, education, roads, sewage, waste disposal) in U. S. cities is lower when the ethnic fragmentation of the city increases.

More recently, Tabellini (2019) finds that the migration of blacks from southern cities to northern cities during the Great Migration had a negative impact on both public spending and tax revenues in northern cities, driven by a reduction in property values. However, while changes in the level of public spending do not necessarily translate into changes in the composition of spending, natives may be less willing to redistribute when facing an increase in ethnic diversity. This could potentially reduce the supply of local public goods (Dahlberg *et al.*, 2012; Munshi & Rosenzweig, 2018). Our work adds to this literature by analyzing the short-term effects of a large international migration event on the composition of local government expenditures, welfare take-up, and changes in property values and rents.

The remainder of the paper is organized as follows: Section 2 presents the theoretical framework. Section 3 presents the data and explains the context. Section 4 presents the empirical strategy and discusses the identification assumptions. Section 5 presents the reduced-form evidence of the effect of immigrant inflows on natives' fiscal contributions, including the robustness checks. Section 6 examines the mechanisms that explain our main findings. Section 7 discusses the implications of our results to the understanding of the overall fiscal effects of immigration. Section 8 concludes.

1.2 *Theoretical Framework*

We begin by laying out a theoretical framework to interpret the empirical results that follow. We present a simple model of equilibrium aggregate tax contributions and government services consumed by natives and immigrants. The starting point is the static framework in Preston (2014).

Suppose the economy consists of an endogenous measure L of workers whom each pay a tax $T(w)$ and consume a range of $k \in \{1, \dots, K\}$ government services G_k . Wages (w) are determined by equating aggregate demand for labor l from perfectly competitive firms with aggregate supply L . We don't model unemployment. Instead, we assume all non-hired labor exit the economy.

Denote the budget cost of providing the k public service by $E_k(G_k, L)$.² This depends both on the nature of the service and size of the population. If the government service is a pure public good, then $\partial E_k / \partial L = 0$, and so its provision is independent of population size. If the service has either the characteristics of a private good or shows congestion in consumption, then $\partial E_k / \partial L > 0$. However, in the case of congested public goods, as opposed to private goods where the cost increases proportionately with the number of consumers, at low levels of consumption, the marginal cost of provision may be below the average cost, but at high levels of consumption (high

² We include debt service as part of government expenditures. Therefore, debt financing of expenditures will affect contributions.

population size) the cost of provision increases more than proportionately with population.

Suppose the population comprises both native workers (N) and immigrant labor (M). Denote the fraction of immigrants as $\theta = M/L$ so that overall population size can be written as $L = N/(1 - \theta)$. The government budget balances if

$$NT(w) + MT(w) - \sum_k E_k(G_k, L) = 0 \quad (1)$$

and natives' net fiscal contributions can be written as

$$NT(w) - \sum_k s_k(\theta) \cdot E_k(G_k, L) = D, \quad (2)$$

where D is a resulting constant and $s_k(\theta)$ is the share of the k public service used by natives³

Assume now that the cost of provision of the k th public service is a function of the mean service use such that $G_k = (1 - \theta)G_k^N + \theta G_k^M$, where G_k^N and G_k^M are the average service use by natives and immigrants, respectively. Differentiating Eq. (2) with respect to the fraction of immigrants (θ) shows that immigration is fiscally beneficial for natives if

$$N \frac{\partial T}{\partial w} \frac{\partial w}{\partial \theta} - \sum_k \left[\frac{\partial s_k}{\partial \theta} E_k(G_k, L) + s_k \left(\frac{\partial E_k}{\partial G_k} \frac{\partial G_k}{\partial \theta} + \frac{\partial E_k}{\partial L} \frac{\partial L}{\partial \theta} \right) \right] < 0. \quad (3)$$

³ Expenditure shares reflect workers' preferences for public goods. We will allow these to also incorporate differences in need-base use of public services (e.g., education or health).

It is fiscally beneficial in the sense that natives would want to pay lower taxes while consuming the same level (or more) of public services⁴. However, because shifts in tax contributions reflect changes in labor market outcomes, lower wages would make natives worse off. A pure estimate of tax payments per capita would imply a benefit from population growth.

Now, if one were to care only about the (net) fiscal comparison between natives and immigrants, where higher (net) fiscal results are preferred ($\partial D/\partial\theta > 0$ in Eq. 3), then immigration would improve natives' fiscal position in the sense that either changes in natives' tax contributions compensate for higher costs of providing public services to a more diverse population or by taxes falling less than the spread in costs from an increase in population size. This can be seen more clearly by setting $\theta = 0$ and working around terms. We can now express natives' fiscal impact from an increase in immigration as:

$$\begin{aligned}
 & \underbrace{\left[\frac{\partial T}{\partial w} \frac{\partial w}{\partial \theta} - T \right]}_{\text{factor price effect}} + \underbrace{\sum_k \left[\frac{E_k(G_k, L)}{L} - \frac{\partial E_k}{\partial L} \right]}_{\text{cost effect}} \\
 & + \underbrace{\frac{1}{N} \sum_k \left[\frac{G_k^M}{G_k^N} E_k(G_k, L) - \frac{\partial E_k}{\partial G_k} (G_k^M - G_k^N) \right]}_{\text{composition and service use effect}} \leq 0.
 \end{aligned} \tag{4}$$

⁴ One can think of θ as capturing an increase in M holding N constant. We abstract from labor displacement of natives to simplify the analysis. As we will show, immigration had no effect on natives' employment. While natives and immigrants compete to some extent for jobs in the informal sector, natives employed in the formal sector are somewhat less exposed to that competition as most immigrants were not allowed to work.

This expression depends on three terms:

- (i) *Factor price effect.* This captures the change in taxes paid by natives from an increase in immigration. In this simplified version, the size and sign of this term will depend on the average wage effect. However, one could expect the overall effect to be different based on the distribution of the native population along the wage distribution. More broadly, this term captures labor market equilibrium effects, both at the extensive and intensive margins, and changes in revenue from capital income.
- (ii) *Cost effect.* This captures the change in the cost of government service provision from a change in the population size. Therefore, the direction of the effect depends on the combination of government services. Note that in the case of public goods—where the level of overall provision of public services is usually not affected by population growth, immigration is beneficial for natives by distributing the cost among a larger population.
- (iii) *Composition and service use effect.* This term captures changes in the composition of government expenditures and differences in the demand for government services between natives and immigrants.

1.3 *Data and Background*

1.3.1 *Data*

Our main data comes from Colombia’s Labor Force Survey (*Gran Encuesta Integrada de Hogares—GEIH*) for 2013–2018. The GEIH is a nationally representative household survey that collects information on labor market conditions. The survey is also representative for the 23 main metropolitan areas (MSAs), which account for up to 50% of the population and 65% of all immigrants. This survey has two crucial advantages. First, it contains information on the geographic distribution of Venezuelan immigrants across MSAs, allowing us to use empirical specifications that rely on variation across geographic units. Second, immigrants are surveyed regardless of their migratory status (regular or irregular).⁵ We define *Venezuelan immigrants* as all individuals born in Venezuela plus all individuals born in Colombia who lived in Venezuela for 1 or 5 years before being surveyed.⁶ We refer to the latter group as native-born returnees or just *returnees*. Therefore, throughout the document, *natives* correspond to all native-born, excluding returnees.

⁵ We follow the terminology used by the Colombian government to classify immigrants. Undocumented or irregular immigrants do not satisfy the requirements established by the host country to enter or remain in the country.

⁶ As is standard in the literature on the fiscal effects of immigration, we classify all native-born dependents (*e.g.*, children aged 18 or less) as immigrants if the head of the household (*e.g.*, parent) is classified as an immigrant. Because we restrict our sample to all those aged 15 to 64 and since our immigration shock is quite recent, excluding native-born dependents does not affect our estimates.

We use the self-reported information in the GEIH to estimate tax contributions to and benefits received from the Central Government, the social security sector, and local governments for all individuals aged 15 to 64. The raw data on all revenues and expenditures of the General Government between 2013 and 2018 comes from Mesa-Guerra & Ramírez-Tobón (2022).⁷ This data includes information on all levels of government for 14 revenue groups (*e.g.*, income tax, wealth tax, indirect taxes, property tax) and 15 expenditure groups (*e.g.*, ‘pure’ public goods, health services, compulsory education, social protection). Section B in the Appendix describes in detail how we estimate the fiscal contribution for each individual observation in the GEIH. We now provide a summary.

Revenues. To estimate personal income tax, we apply year-specific tax rates to reported income, including capital income. Social insurance and payroll taxes are estimated by applying the tax schedule based on an individual’s affiliation status to a pension fund and a health regime. Contributions from VAT and other indirect taxes are estimated using decile-specific effective tax rates from Mesa-Guerra & Ramírez-Tobón (2022) and applying them to gross household income. Property tax and wealth tax contributions are estimated using proxies for asset ownership and estimates of property values. Other contributions such as motor vehicle, industry, and phone taxes are allocated using information on

⁷ The General Government consist of the National government, regional and local governments, and the social security sector.

vehicle ownership (car or motorcycle), registered businesses, or having a landline. In the case of the financial transaction tax, we apply estimates of the expenditures-to-income ratio by decile to individual income reported in the GEIH and use the tax exemption threshold to allocate revenues. All other revenue sources are allocated per capita using eligibility information in the GEIH, as described in detail in the Appendix.

Expenditures. To allocate government expenditures, we followed a simple rule: all self-reported income or subsidies from government agencies are taken as such, and the rest are assigned per capita using individual eligibility. In the first group lies all spending on welfare or social protection programs: sickness and disability, old age, family and children, unemployment, housing, and vulnerable population. The second group comprises public goods (rival and non-rival in consumption), law courts and prisons, water supply, health, education, and debt service. Public goods are assigned per capita; law and prison spending is assigned using the share of the prison population by origin country; water spending is allocated using the population with access to the water supply system; health services are allocated using the individual health cost by age group for each regime; education services are allocated based on enrollment, differentiating between compulsory education, job training, and higher education; debt service is allocated per capita but conditioning on the year of arrival to the country.

Table 1.1 shows descriptive statistics for natives and Venezuelan immigrants aged 15 to 64, divided by group. It stands out the fact that Venezuelan immigrants went from representing around 0.2% of the population in 2013 to representing 3.1% of the population by 2018. A first look at the demographic characteristics of natives and Venezuelan immigrants aged 15 to 64 suggests that the latter are younger and have more years of schooling. However, this is true for Venezuelan-born immigrants who are, on average, seven years younger and have an additional year of education relative to natives. In contrast, as of 2018, Colombian-born returnees were slightly older and had lower schooling than natives.

Regarding labor market conditions, Venezuelan immigrants contributed substantially to the overall increase in the labor force between 2013 and 2018. Immigrants have both a higher employment and unemployment rate compared to natives. As rows 8 to 10 suggest, immigrants are more likely to be self-employed and potentially downgrading upon their arrival.

Figure 1.16 in the Appendix shows the scope of Venezuelans' downgrading upon arrival. If Venezuelan-born immigrants received the same returns to education and potential experience as natives (red line), they should be less overrepresented in the lower part of natives' wage distribution (below the 35th percentile) and more represented in the upper

part of the wage distribution (over the 60th percentile). Therefore, the decrease in the average monthly labor income and the increase in the share of workers with earnings below the legal minimum wage observed in rows 9 and 10 of Table 1.1 is consistent with a downgrading story.

Table 1.1 Descriptive statistics for 2013 and 2018

Characteristics	Natives		Venezuelan immigration			
	2013	2018	<i>Returnees</i>		<i>Venezuelan-born</i>	
			2013	2018	2013	2018
<i>Demographics</i>						
1. Share of the population	99.8	96.9	0.1	0.9	0.1	2.2
2. Percent Male	48.7	48.8	55.2	51.6	48.3	50.8
3. Average Age (years)	35.7	36.3	34.2	37.5	31.2	29.4
4. Avg. years of schooling (age 15+)	9.0	9.5	8.9	7.9	10.5	10.5
<i>Labor Market</i>						
5. Share of the labor force	99.7	96.5	0.2	1.0	0.1	2.5
6. Employment Rate	66.5	66.3	72.3	70.7	66.3	70.0
7. Unemployment Rate	9.9	9.9	12.7	12.2	14.0	14.6
8. Percent Self-Employed	46.2	45.9	52.3	55.0	43.2	48.8
9. Avg. Monthly Labor Income (K)	998.4	1,023.4	870.2	678.1	2,479.8	740.3
10. Percent Earning below min. wage	45.5	44.1	47.4	61.5	40.0	63.5

Notes: The table reports for 2013 and 2018 descriptive statistics for natives, Venezuelan-born immigrants, and returnees aged 15 to 64. The employment rate is defined as the ratio of the employed to the working-age population. Average real labor income (adjusted using the 2018 Consumer Price Index as base year) includes all labor income for wage and salary workers and self-employed, excluding business owners. *Source:* Own estimates using data from the GEIH.

1.3.2 Background

According to the Interagency Coordination Platform for Refugees and Migrants (R4V), there are more than 7 million Venezuelan immigrants around the world. Most of this population emigrated after 2014 following the social, economic, and political crisis that has plunged the country into

poverty.⁸ Over 80% of all outflows from Venezuela have found refuge in Latin American countries. Colombia has been the primary recipient of immigrants. Between 2013 and 2018, the total stock of Venezuelan immigrants in Colombia (including returnees) went from 0.1 to 1.5 million, a fifteen-fold increase. For that same period, the stock of those aged 15 to 64 increased by 925 thousand. As of 2018, around 63% of the working-age immigrants from Venezuela were living in urban areas (see Figure 1.1). Immigrants aged 15 to 64 explained 67% of the overall increase in Venezuelan immigrants. To get a sense of the magnitude of the supply shock, immigrants from Venezuela went from representing only 0.3% of the native working-age population in 2013 to 3.8% in 2018.

However, Venezuela's economic and political crisis also forced many Colombians to return to their country. Inflows of returnees were particularly intense between August 2015 and August 2016 when the Venezuelan government unilaterally decided to close the border between the two countries and expel many Colombians living in the border regions in Venezuela. More than 23,000 Colombians were either expelled or voluntarily returned to Colombia during that period. With more than 1,700

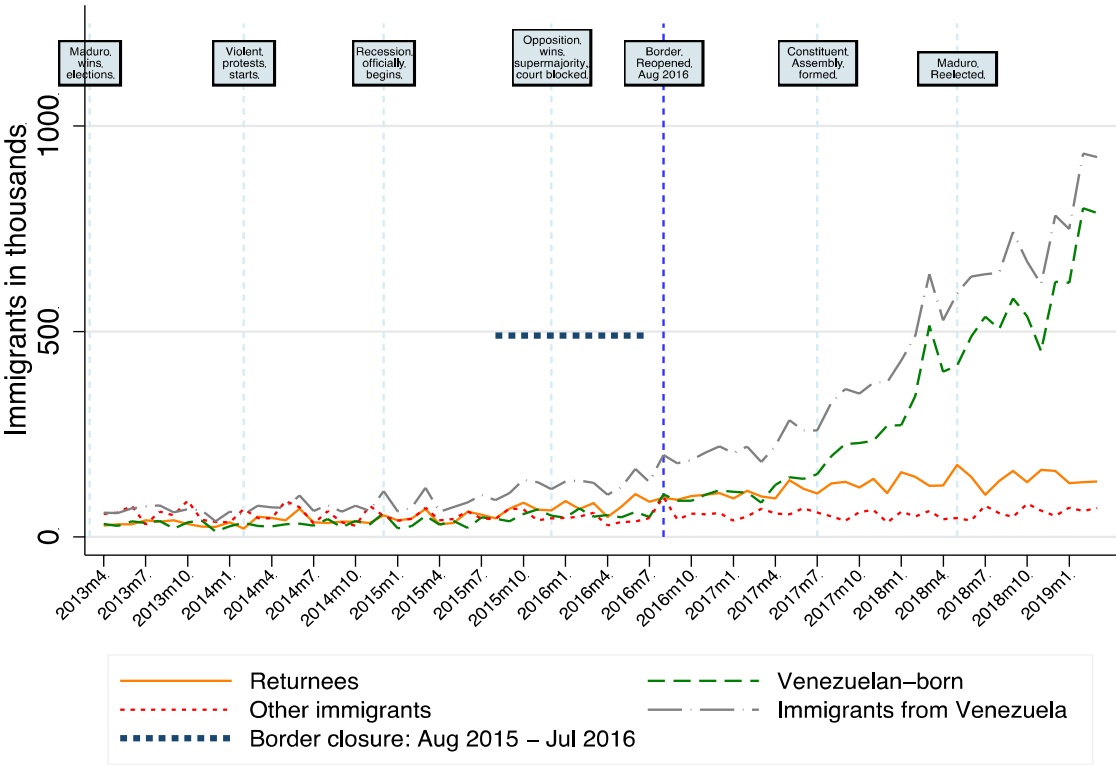
⁸ In a span of 5 years (2013 to 2018), the Venezuelan economy contracted by half. According to the 2021 ENCOVI survey (*Encuesta Sobre Condiciones de Vida en Venezuela*), conducted by three private universities in Venezuela, 94.5% of the population was living in poverty and 76.6% was living in extreme poverty.

being deported.⁹ As shown in Figure 1.1, after the border was reopened on August 12, 2016, immigration from Venezuela dramatically increased. For our estimations, we treat 2013 and 2014 as pre-shock periods as we observed very low inflows from Venezuela.

Following the significant increase in immigrant inflows, at the beginning of 2017, the government created a two-year special permit (*Permiso Especial de Permanencia—PEP*) that allowed Venezuelan-born immigrants with regular status to stay and work in the country. It was estimated that two out of every three Venezuelan-born immigrants were under an irregular status (Reina *et al.*, 2018), either because they overstayed in the country or because they entered through unauthorized border crossings. As a result, the Colombian government expanded this program in 2018 to cover around 440,000 undocumented immigrants that had voluntarily registered at the time using the Administrative Register of Migrants from Venezuela (RAMV). This legal mechanism was intended to improve the assimilation of this population, allowing immigrants to work and access health and education services.

⁹https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/151015_informe_de_situacion_no_12_situacion_de_frontera_final.pdf

Figure 1.1 Migrant stock, April 2013 - March 2019



Notes: The Figure depicts the evolution of the monthly stock (in thousands) of working-age (15-64 years old) immigrants from Venezuela and other immigrant groups living in urban areas between April 2013 and March 2019. The gray dashed line corresponds to the sum of all migrants coming from Venezuela, including returnees and Venezuelan-born. The horizontal blue dashed line indicates the moment the Venezuelan government unilaterally decided to close the main border crossings between the two countries. The vertical dashed blue line indicates the date when the border was reopened. Light-blue boxes on the upper part of the graph show important political and economic events in Venezuela. Returnees are defined as Colombian-born individuals returning from Venezuela. *Source:* Own estimates using data from the GEIH 2013-2019.

1.4 Empirical Framework

1.4.1 Estimation Strategy

Considering our theoretical framework, an analysis of the impact of immigration on natives’ fiscal position should relate the changes in natives’ net fiscal contributions over time to the corresponding change in population due to immigration (see Card & Peri, 2016). As immigration is

distributed across local labor markets and a large share of government finances is accrued at the regional and local level, one can use the variation across geographical units. Let M_{jt} and N_{jt} represent the number of immigrants and natives in area j in year t , respectively. Then, the empirical specification that derives directly from the relationship of interest is:

$$\Delta y_{jt} = \beta \Delta m_{jt} + \Omega_t + \varepsilon_{jt}, \quad (5)$$

where

$$\Delta m_{jt} = \frac{M_{jt} - M_{jt-1}}{M_{jt-1} + N_{jt-1}} = \frac{\Delta M_{jt}}{L_{jt-1}}, \quad (6)$$

and Δy_{jt} is the residual change in the average net fiscal contributions of natives. We also include year-fixed effects (Ω_t) and control for changes in the demographic and skill composition of the native population. Instead of relying on averages at the aggregate level, we account for changes in demographic characteristics by exploiting individual-level data and estimating fiscal contributions net of individual characteristics. For each year between 2013 and 2018, we regress individual fiscal contributions on age, age squared, sex, and education. We then use the resulting residuals to estimate equation (5). Finally, ε_{jt} is a random error term. We restrict our sample to all individuals aged 15 to 64.¹⁰

¹⁰ To alleviate the potential effect of outliers, specially coming from the allocation of corporate taxes to resident households which were allocated equally to those that reported

We allow cities to be on differential trends by interacting year dummies with a variety of 2013 city characteristics such as the share of local revenues that accrue to transfers from the central government, the share of spending on public goods, the share of working age population in the city, the share of college workers, and the share of workers employed in manufacturing. These are meant to pick up any potential correlation between natives' change in fiscal contributions and a city's fiscal or demographic composition. By fixing controls in a baseline year and not conditioning on variables measured in the post-treatment period, we avoid potential pitfalls from controlling for intermediate outcomes that may have been affected by the shock.

Since the economy also experienced a shock from the fall in commodity prices starting in late 2014, which considerably affected the fiscal balance of the General Government, by including year-fixed effects, we control for common changes that may have affected local budgets. Figure 1.8 in the Appendix shows the evolution of average fiscal contributions compared to the immigrant shock (Panel A) and the cyclical fiscal balance for the energy sector (Panel B).¹¹ While fiscal contributions

to have received individual dividend or interest income, our main results are computed by trimming the distribution of net fiscal contributions each year for observations below the 1st percentile and above the 99th percentile. Corporate and capital taxes assigned to resident households' range between 25% and 34% of total government revenues between 2013 and 2018.

¹¹ The energy cyclical balance measures the change in fiscal revenues caused by the difference between the observed price and the long-term price of crude oil of the previous period.

seem to follow the energy cycle, because this is a common shock, it will not affect our results.

1.4.2 Identification

If the fraction of immigrants in the local labor market population is correlated with local demand shocks, the coefficient of interest (β) will be biased. Given that the location of immigrants is the result of decisions that depend on conditions at the destination, a simple comparison between high- and low-immigration areas would result in biased estimates. In other words, if immigrants may endogenously select themselves into areas experiencing positive demand shocks, this would cause β in equation (5) to be upward biased. In addition, because labor force surveys are not initially intended to be a representative sample of all immigrant populations but of the overall population, this could lead to measurement error and, therefore, attenuation bias (Aydemir & Borjas, 2011).

We follow the empirical literature and estimate Δm_{jt} using historical settlement patterns of Venezuelan immigrants in Colombia. The instrument takes advantage of the well-known fact that new immigrants tend to locate near earlier immigrants from the same origin country, creating a source of exogenous variation (Altonji & Card, 1991; Card, 2001). We use census data to construct this instrument of predicted inflows to each area as a fraction of the corresponding local population in 2013:

$$IV_{jt}^{Past} = \left(\frac{M_{j,2005}}{M_{2005}} \right) \times \left(\frac{1}{L_{j,2013}} \right) \times \Delta M_t. \quad (7)$$

Here, M_{2005} is the number of Venezuelan-born immigrants in Colombia in 2005, $M_{j,2005}$ is the number of Venezuelan-born immigrants in location j in 2005, $L_{j,2013}$ is the working-age population of location j in 2013, and ΔM_t is the number of working-age individuals arriving in Colombia each year from Venezuela.¹² Notice that the share $\left(\frac{M_{j,2005}}{M_{2005}} \right)$ includes all Venezuelan-born immigrants in Colombia, while the shift (ΔM_t), restricts inflows to all working-age Venezuelans arriving to the country.

Identification comes from exogenous exposure (given by the shares) to a common shock (given by the shift), as discussed by Goldsmith-Pinkham *et al.* (2020). The validity of the instrument depends on the exogenous variation in the national inflow rates from Venezuelan to local conditions in a specific city. Because increasing inflows of Venezuelan immigrants during our period of analysis are driven primarily by push factors in Venezuela that are uncorrelated with specific conditions in Colombia, this condition is very likely to hold.¹³ Goldsmith-Pinkham *et al.*

¹² Inflows from Venezuela include individuals classified as Venezuelan-born, native-born returnees, and other foreign-born who previously lived in Venezuela. For 2005 we use total counts derived directly from the complete Census as published online by DANE. For 1973 and 1993 we use the 10% census samples from IMPUS.

¹³ A fact supporting this assumption is that immigration from other countries did not significantly change between 2013 and 2018 as depicted in Figure 1.

(2020) show that if the pre-period shares are correlated with unobserved local conditions, even if the national inflow rates are not correlated to those conditions, the instrument might not satisfy the exclusion restriction. Following equations (5) and (7), our empirical strategy tests whether differential exposure to the migration shock leads to differential changes in fiscal contributions for working-age natives. By focusing on changes, we can control for unobserved level differences. As long as there are no time-varying omitted local characteristics that are correlated both with our instrument and the outcome variable, our empirical strategy is thus valid.

The directed acyclic graph (DAG) in Figure 1.9 in the Appendix illustrates our setup, where the reduced-form relationship between natives' fiscal contributions and immigrant inflows is confounded by unobserved local demand factors, thus the need for an instrument. In addition, it highlights some of our assumptions. First, the only way in which immigration influences natives' fiscal contributions is through changes in factor prices (*e.g.*, labor market effects), the cost of providing public services, or the allocation of public spending, as presented in Eq. (4). Second, because the relationship between immigration and the mediating channels is confounded by unobserved local factors, this can be uncovered by ways of an instrument.

To build credibility for our empirical design, we follow the three-step approach suggested by Goldsmith-Pinkham *et al.* (2020). First, we look at the correlation between the exposure shares and local characteristics in 2005 and 2013. Table 1.2 indicates that local characteristics explain a good amount of the cross-sectional variation in the shares, with over half of the variation being explained when accounting for observed characteristics in 2013. In addition, we do not find a significant relationship between our local characteristics and the exposure shares or the instrument, except for self-employment in 2013. Although the share of self-employed workers in 2013 is not a mediator of the effect of immigration on natives' fiscal contributions (see Figure 1.9 in the Appendix), including this as a control in Eq. (5) would potentially bias our estimates. First, it would be equivalent to partially controlling for labor market outcomes. Second, the fact that the share of self-employed is a strong predictor of the distribution of immigrants would amplify the bias created by the unobserved local demand factors (Pearl, 2013). The results presented in Table 1.2 suggest no other channels through which the shares affect outcomes in the relevant period except by predicting where the new inflows of immigrants locate within the country.

Table 1.2 Relationship between exposure shares and local characteristics

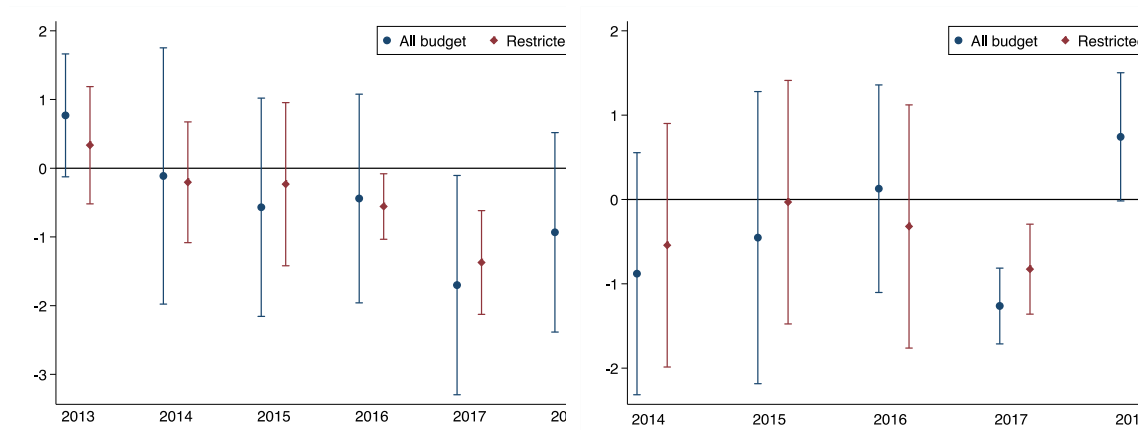
Characteristics	<i>Exposure share</i>				<i>Instrument</i>	
	Coef.	SE	Coef.	SE	Coef.	SE
<i>Panel A. Local characteristics in 2005</i>						
Share of population age 15 to 64	0.596	(0.717)				
Share of college-educated population	0.240	(0.375)				
Share of workers self-employed	-0.397	(0.489)				
Share of employment in manufacturing	-0.848	(0.711)				
<i>Panel B. Local characteristics in 2013</i>						
Share of transfers from the central gov.			-0.279	(0.15)	-0.015	(0.19)
Share of expenditures in public goods			0.110	(0.08)	0.120	(0.11)
Log labor income			0.114	(0.06)	0.044	(0.08)
Share of population age 15 to 64			-0.755	(0.67)	0.020	(0.89)
Share of college-educated population			0.369	(0.46)	-0.153	(0.60)
Share of workers self-employed			0.516	(0.19)	0.742	(0.25)
Share of employment in manufacturing			0.506	(0.46)	0.480	(0.60)
Obs.	23		23		23	
R^2	0.365		0.669		0.550	

Notes: Each column reports the results of a single regression of the 2005 Venezuelan share on a set of local characteristics in 2005 and 2013. The final column uses as dependent variable the immigrant enclave instrument constructed using the change in immigrant inflows from 2013 to 2018. Results are weighted by the 2013 population. Standard errors in parentheses.

As a second step, we test for parallel pre-trends for our baseline period 2013-2014 (*differences-in-differences* logic), following our discussion in Section 3.2. We construct our pre-trend figures by estimating for each year a reduced-form regression of natives' fiscal contributions on the share of Venezuelan immigrants in 2005 and including the set of controls in Table 1.2, panel B. We do this for both levels and changes in the dependent variable. Figure 1.2 shows that the variation in the share of Venezuelan immigrants in 2005 did not predict lower net fiscal contributions for natives between 2013 and 2014, conditional on local characteristics. The figures suggest that a large enough change in the stock of immigrants starting in 2015 led to lower fiscal contributions for working-age natives.

As a third step, in Section 5.2. we consider an alternative instrument to corroborate our main results.

Figure 1.2 Pre-trends using the exposure shares in 2005
 Panel A. Levels Panel B. Changes



Notes: The figures report year-by-year estimates of a reduced-form regression of natives' fiscal contributions on the share of Venezuelan immigrants in 2005. We present pre-trend results for the dependent variable both in levels (Panel A) and first differences (Panel B). Regressions include the set of controls described in Table 1.2, panel B. Results are computed by trimming the distribution of net fiscal contributions each year at 1% and 99%. Estimates are weighted by the working-age population in 2013.

An issue raised by Jaeger *et al.* (2019) is that when using immigrant enclave instruments, if local characteristics at the base period influence future levels of migration inflows and thus persist over time, the coefficient of interest will be biased. This could be particularly relevant when estimating effects on labor market outcomes as settlement patterns of immigrants are persistent. For example, estimates of the impact of immigration on wages may confound the partial equilibrium wage impact

of recent immigration inflows (presumably negative), with the local labor adjustment to previous immigrant supply shocks (presumably positive), leading to a ‘dynamic adjustment bias. Jaeger *et al.* (2019) show that the dynamic adjustment bias is reduced in settings in which the overall rate of immigration has temporarily increased and where origin-specific push factors change the migration rate of a particular origin country. Our empirical setting closely resembles these conditions. In addition, the literature has shown, at least in the developed world, that the effects of previous immigrant supply shocks tend to die out within ten years (National Academies of Sciences, Engineering, and Medicine, 2017). In Section 5.2. we show estimates allowing for dynamic adjustment to corroborate our findings.

Finally, instrumental variables could also account for the attenuation bias generated by the low sampling of Venezuelan immigrants observed in the GEIH in its early years. This will be the case if the measurement error in the selected instrument is uncorrelated with the measurement error in the spatial distribution of Venezuelans for the period 2013–2018. Since our instruments are constructed using counts of immigrants from the 2005 population census (and from the 1993 census in some robustness exercises), the presence of measurement error in the instruments is not a threat to the consistency of our IV approach.

1.5 Results

1.5.1 Effect on Natives' Fiscal Contributions

We begin by estimating the effect of changes in relative immigrant inflows on the change in natives' fiscal contributions net of individual characteristics, following Eq. (5). In Table 1.3, we report the results of the change in immigrant inflows expressed as a fraction of the size of the total local population in the previous year. In addition, we present our results using the complete fiscal contributions [Panel (a)] and using a restricted budget that excludes revenues and expenditures that are not derived from individual behavior [Panel (b)]. In the latter, we exclude corporate taxes, gross operating surplus, royalties, 'pure' public goods, and debt service.

Columns 1–3 present the estimates using an ordinary least squares (OLS) estimator, while columns 4–6 use a two-stage least square (2SLS) estimator based on the immigrant enclave instrument from equation (7). Columns 2 and 5 include only pretreatment covariates, while columns 3 and 6 include pretreatment covariates interacted with year dummies to allow cities to be on differential trends. As described above, we exclude from our set of controls the share of self-employed workers as this variable is highly correlated with the exposure shares and will thus introduce bias to our estimates.

Table 1.3 Effect of immigration on net fiscal contributions – first differences
(COP\$ million, 2018 equivalent)

Explanatory variable	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel (a): Contributions using the complete budget</i>						
Immigrant inflows (m_{jt})	-0.009 (0.027)	-0.008 (0.032)	-0.022 (0.029)	-0.006 (0.031)	-0.000 (0.035)	0.007 (0.036)
AR test (<i>p-value</i>)				0.860	0.99	0.85
<i>Panel (b): Contributions using the restricted budget</i>						
Immigrant inflows (m_{jt})	0.002 (0.026)	0.008 (0.029)	-0.005 (0.028)	-0.018 (0.024)	-0.019 (0.03)	-0.000 (0.03)
AR test (<i>p-value</i>)				0.441	0.529	0.988
Kleibergen-Paap <i>F</i> -stat				13.84	17.34	24.21
Year FE	✓	✓	✓	✓	✓	✓
MSA-level controls		✓			✓	
Controls x Year FE			✓			✓
Obs.	115	115	115	115	115	115

Notes: The table reports the coefficients obtained by regressing the change in natives' net fiscal contributions net of individual characteristics on the change in immigrant inflows (Δm_{jt}) between 2013–2018 and year fixed effects. Individual characteristics include a dummy for sex, age, and education. Education groups are classified as: (i) less than HS, (ii) HS graduate, (iii) some college, (iv) bachelor's degree, (v) any post-bachelors. Residuals for observation with no information on education are obtained from conditioning only on sex and age. MSA-level controls include the share of local revenues that accrue to transfers from the central government, the share of expenditures in public goods, the share of working age population in the city, the share of college workers, and the share of workers employed in manufacturing. We report the Anderson–Rubin test of structural parameters that is robust to the presence of weak instruments. AR test the hypothesis $H_0: \beta = 0$ by estimating a reduced form equation for the dependent variable with the full set of instruments as regressors and testing that the coefficients of the excluded instruments are jointly equal to zero. Results are computed by trimming the distribution of net fiscal contributions each year at 1% and 99%. Estimates are weighted by the working-age population in 2013. Robust standard errors are reported in parentheses. *** Denotes significance at 1%, ** significance at 5% and * significance at 10%.

Overall, we find that the effect of immigration on natives' average net fiscal contributions is small and not distinct from zero. This is the case when using either the complete budget or our restricted definition. Looking at OLS results in columns 1–3, if immigrants were positively sorting into

high-wage areas, one would expect larger estimates (in absolute terms) than those produced from 2SLS regressions. As we mentioned before, the potential for measurement error would lead to attenuation bias in the estimated effect under OLS, which we expect to correct with the inclusion of our instrument.

As suggested by Andrews *et al.* (2019), we report the Anderson–Rubin (AR) test of structural parameters that is robust to the presence of weak instruments. Based on the AR test, identification-robust confidence intervals would include zero at any standard confidence level.

The recurring negative coefficients presented in Table 1.3 and the fact that the pace of immigrant inflows increased exponentially from 2017 indicates that it is likely that we see effects at the end of our sampled period. To corroborate this intuition, we present year-by-year estimates using Eq. (5) in Figure 1.10 in the Appendix. Panel A suggests a small negative effect in 2017. According to these estimates, a 1 percentage point increase in the inflow of immigrants relative to the working-age population in the previous year would lead to a decrease in average net fiscal contributions for natives of about COP\$0.1 million, but with no effect on the following year.

Considering that the average increase in the size of the working-age immigrant population over the period studied was around 0.72 percentage points per year, the actual impact is close to COP\$0.07 million. This is

equivalent to about \$23 US dollars in 2018, or 0.3% of GDP per capita. Not a large effect. We also present estimates of natives' change in fiscal contributions relative to our baseline year (2013) in Figure 1.11 in the Appendix. Results in Panel A presents evidence of significant negative effects starting in 2015 that decline over the following years.

We go further and estimate the effect separately for tax contributions and benefits received in the form of expenditures. Results displayed in Table 1.4 suggest no significant effect on average tax contributions and a positive effect on expenditures only when we allow for differential trends in fiscal contributions across areas. Based on the identification-robust AR test, the effect is not distinguishable from zero at the 5% level.

1.5.2 Robustness Checks

To assess the robustness of our estimates, we conduct additional exercises using alternative regression specifications. In Table 1.5, we report results using the restricted budget definition. Table 1.7 in the Appendix presents the results using the complete budget. We report all coefficients with their corresponding standard errors, the first-stage F -statistic, and identification-robust Anderson–Rubin confidence intervals.

In row (a), we use the shares from the 1993 census. In row (b), we modify Eq. (5) and normalize the fraction of immigrants from Venezuela in

each area by the local population in 2013.¹⁴ By fixing the population in a base year, we intend to avoid potential spurious correlation between fiscal contributions and changes in the fraction of immigrants that may have been induced by changes in the native population—which may carry over with the lagged terms.

Table 1.4 Effect of immigration on revenues and expenditures
(COP\$ million, 2018 equivalent)

Explanatory variable	Revenues			Expenditures		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel (a): Contributions using the complete budget</i>						
Immigrant inflows (m_{jt})	-0.020 (0.039)	-0.012 (0.042)	0.069 (0.042)	-0.016 (0.026)	-0.017 (0.029)	0.049* (0.026)
AR test (<i>p-value</i>)	0.590	0.778	0.051	0.510	0.521	0.064
<i>Panel (b): Contributions using the restricted budget</i>						
Immigrant inflows (m_{jt})	-0.045 (0.045)	-0.053 (0.047)	0.045 (0.043)	-0.023 (0.026)	-0.031 (0.027)	0.038* (0.021)
AR test (<i>p-value</i>)	0.267	0.224	0.268	0.300	0.189	0.069
Kleibergen-Paap <i>F</i> -stat	13.841	17.343	24.212	13.841	17.343	24.212
Year FE	✓	✓	✓	✓	✓	✓
MSA-level controls		✓			✓	
Controls x Year FE			✓			✓
Obs.	115	115	115	115	115	115

Notes: The table reports the coefficients obtained by regressing the change in natives' tax contributions and benefits net of individual characteristics on the change in immigrant inflows (Δm_{jt}) between 2013–2018 and year fixed effects. Individual characteristics include a dummy for sex, age, and education. Education groups are classified as: (i) less than HS, (ii) HS graduate, (iii) some college, (iv) bachelor's degree, (v) any post-bachelors. Residuals for observation with no information on education are obtained from conditioning only on sex and age. MSA-level controls include the share of local revenues that accrue to transfers from the central government, the share of expenditures in public goods, the share of working age population in the city, the share of college workers, and the share of workers employed in manufacturing. We report the Anderson–Rubin test of structural parameters that is robust to the presence of weak instruments. Results are computed by trimming the distribution of net fiscal contributions each year at 1% and 99%. Estimates are weighted by the working-age population in 2013. Robust standard errors are reported in parentheses. *** Denotes significance at 1%, ** significance at 5% and * significance at 10%.

¹⁴ The correlation between $\Delta M_t/L_{t-1}$ and $\Delta M_t/L_{2013}$ is 0.9983.

In row (c), we exploit the microdata from the GEIH and estimate the following modified version of Eq. (5) using individual observations:

$$y_{ijt} = \beta \tilde{m}_{jt} + \gamma \mathbf{X}_{ijt} + \phi \mathbf{Z}_{j2013} \cdot \mathbf{1}[\Omega_t = t] + \Omega_j + \Omega_t + \varepsilon_{ijt}, \quad (8)$$

where the fraction of immigrants is defined as $\tilde{m}_{jt} = M_{jt}/L_{j,2013}$.¹⁵ We also include a vector of individual-specific characteristics, \mathbf{X}_{ijt} , interact time-invariant fiscal controls (\mathbf{Z}_{j2013}) with time fixed effects (Ω_t), and include area fixed effects, Ω_j . Including area-fixed effects is approximately equivalent to estimating a first difference model in which the variable of interest is the change in the fraction of immigrants in a given metropolitan area. Since we fixed the total population in the area in the immigrant ratio to the population level in the baseline period (2013), a first-order approximation of the change in the fraction of immigrants in a given area results in $\Delta M_t/L_{2013}$.¹⁶

In row (d), we construct a distance-based instrument in the spirit of Del Carpio & Wagner (2015) and Delgado-Prieto (2021). The instrument

¹⁵ Following Eq. (7) we construct an enclave instrument using as “shift” the total stock of Venezuelan immigrants in Colombia in each year.

¹⁶ Fixing the denominator to the baseline population eliminates the bias induced by changes in the number of native workers as a result of changes in immigrant inflows. Since we are interested in estimating the effect of immigration on the fiscal outcomes of natives (left-hand side of Eq. 8), changes in the number of native workers that appear on the right-hand side of Eq. (8) can lead to bias in the partial correlation between native outcomes and immigrant inflows as showed by Card & Peri (2016).

predicts the number of immigrants that arrived in each city j in year t as follows:

$$IV_{jt}^{Dist} = \sum_v \frac{1}{T_{vd}} \theta_v \cdot \Delta M_t, \quad (9)$$

where T_{vd} is the road distance between the centroid of the capital city in each Venezuelan province v and the centroid of the capital city in each Colombian department d ; ¹⁷ θ_v is the distribution of Venezuelan-born immigrants by region of origin in Venezuela from the voluntary registry of undocumented migrants (Administrative Register of Migrants from Venezuela—RAMV) conducted in 2018; ¹⁸ ΔM_t corresponds to the new inflows of Venezuelan immigrants to Colombia in year t . Ideally, one would want information on the distribution of Venezuelan immigrants by city/region of origin before the shock. However, the RAMV is the only available source of information that we know of that contains the region of origin of Venezuelan immigrants. We assume that the share of immigrants from each sending region has remained stable over the years as it is a product of information networks that have been built—at least—over the last three decades.

Since the two countries share a border of more than 2,000 km., the instrument is based on the idea that distance creates time and economic

¹⁷ The distance is calculated using the package *georoute* in Stata 17 (Weber & Péclat, 2017).

¹⁸ The information can be accessed here: <https://data.unhcr.org/en/documents/details/64101>

costs for new immigrants, and therefore, it is a crucial determinant of immigrants' location decisions in Colombia.¹⁹ Identification comes from comparing cities in areas close to the border with those further away. A major threat to identification with this instrument is that cities closer to the border could have been more affected by economic shocks arising from changes in trade patterns between Colombia and Venezuela which could be correlated with changes in the economic outcomes of natives and affect their fiscal contributions (Bonilla *et al.*, 2020; Delgado-Prieto, 2021).

To control for a potential dynamic bias, row (e) adds a term for lagged immigrant inflows and instruments with a lagged version of our enclave instrument from Eq. (7), as suggested by Jaeger *et al.* (2019). Finally, row (f) estimates Eq. (5) using a limited information maximum likelihood (LIML) estimator. 2SLS and LIML yield identical estimates of $\hat{\beta}$ in the exactly identified case, but LIML has better small sample properties.

¹⁹ The location of Colombians in Venezuela before the shock also serves an information mechanism that conveys information about where to locate in Colombia.

Table 1.5 Robustness estimates of the effect of immigration on natives' fiscal contributions – restricted budget

<i>Estimates of immigrant inflows (m_{jt})</i>	Coef.	SE	<i>F</i> -stat	Anderson–Rubin CI
<i>(1) Net fiscal contributions</i>				
(a) Using the shares from the 1993 census	0.005	0.028	23.707	[-0.040, 0.078]
(b) Normalizing ΔM_{jt} by the local pop. in 2013	0.000	0.030	22.672	[-0.053, 0.074]
(c) Using individual pooled data (\tilde{m}_{jt})	0.006	0.014	25.710	[-0.022, 0.033]
(d) Using a distance-based instrument	-0.008	0.027	66.698	[-0.056, 0.045]
(e) Controlling for dynamic bias				
– Contemporaneous term: m_{jt}	0.105	0.083	30.433	[-0.069, 0.336]
– Lagged term: m_{jt-1}	-0.157	0.115	74.634	[-0.396, 0.082]
(f) LIML	0.000	0.031	24.212	[-0.047, 0.075]
<i>(2) Revenues</i>				
(a) Using the shares from the 1993 census	0.052	0.038	23.707	[-0.010, 0.153]
(b) Normalizing ΔM_{jt} by the local pop. in 2013	0.044	0.042	22.672	[-0.031, 0.146]
(c) Using individual pooled data (\tilde{m}_{jt})	0.024	0.014	25.825	[-0.007, 0.050]
(d) Using a distance-based instrument	0.026	0.038	66.698	[-0.041, 0.106]
(e) Controlling for dynamic bias				
– Contemporaneous term: m_{jt}	0.132	0.109	30.433	[-0.094, 0.443]
– Lagged term: m_{jt-1}	-0.131	0.143	74.634	[-0.428, 0.166]
(f) LIML	0.045	0.043	24.212	[-0.027, 0.149]
<i>(3) Expenditures</i>				
(a) Using the shares from the 1993 census	0.039	0.019	23.707	[0.005, 0.086]
(b) Normalizing ΔM_{jt} by the local pop. in 2013	0.037	0.020	22.672	[-0.002, 0.083]
(c) Using individual pooled data (\tilde{m}_{jt})	0.017	0.007	25.432	[-0.002, 0.027]
(d) Using a distance-based instrument	0.028	0.018	66.698	[-0.004, 0.067]
(e) Controlling for dynamic bias				
– Contemporaneous term: m_{jt}	0.023	0.050	30.433	[-0.116, 0.127]
– Lagged term: m_{jt-1}	0.023	0.048	74.634	[-0.078, 0.124]
(f) LIML	0.038	0.021	24.212	[-0.002, 0.084]

Notes: The table reports various estimates of the effect of changes in the fraction of immigrants on natives' net fiscal contributions, tax contributions, and expenditures. All regressions include year dummies and interactions of MSA-level controls with year dummies. Results are net of individual-level controls (sex, age, education) and computed by trimming the distribution of contributions each year at 1% and 99%. We report 5%-level identification-robust Anderson–Rubin confidence sets. Results are expressed as 2018 equivalent COP\$ million.

The results presented in Table 1.5 suggest, as stated before, that natives' net fiscal contributions were not affected by immigrant inflows. A key takeaway from the dynamic analysis, even though the results are not distinct from zero, is that contemporaneous inflows have a positive coefficient while lagged inflows have a negative coefficient. This seems to point in the direction of immigrants improving natives' fiscal position by spreading costs in the short-term, but possibly offsetting this with an effect on factor prices in the medium- to long-term. We still do not find any effects on average tax contributions and clear evidence of an increase in average benefits received—except when using the 1993 immigrant shares.

1.5.3 Differences by Revenue and Expenditure Group

We explore further the effects on natives' fiscal contributions by looking at changes across revenue and expenditure groups. Figure 1.3 presents a detailed picture in each case by plotting the point estimates and the respective confidence intervals of our OLS and 2SLS coefficients obtained by regressing natives' individual contributions on the fraction of immigrants, conditioning on baseline city characteristics. In the Appendix, we provide a complete description of how we defined and computed each revenue and expenditure category.

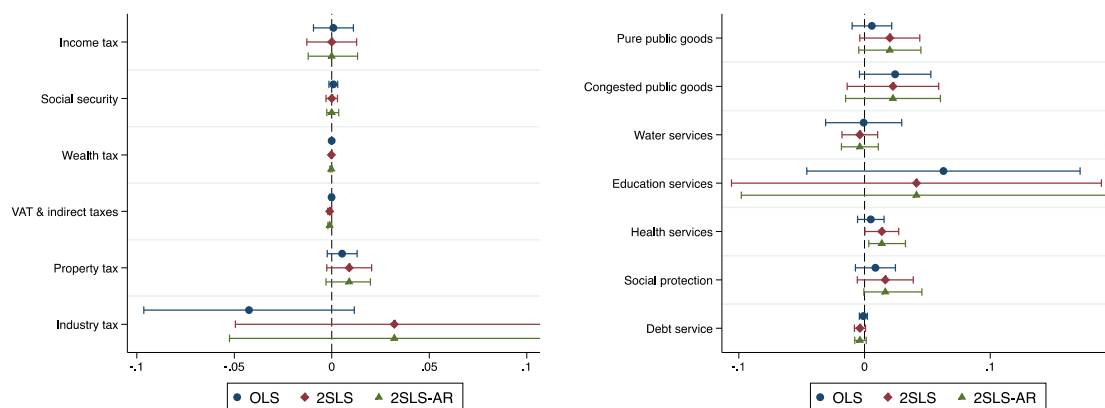
Results presented in Figure 1.3 suggest that changes in immigrant inflows over the period had only a small negative effect on indirect taxes on

the revenues side and a positive effect on health services in the case of expenditures. These findings have two implications. First, since we are not seeing effects on personal income tax and social security contributions, changes in indirect taxes seem to be driven primarily by changes in household income. Second, a positive effect on health services could reflect one or both of the following: an increase in the unit cost of health services from changes in the epidemiological and demographic profile of the population or an increase in coverage of the immigrant population through demand subsidies which are then borne to a certain extent by the native population. We explore these different mechanisms in detail next.

Figure 1.3 Change in natives' fiscal contributions by revenue and expenditure group (*COP\$ million, 2018 equivalent*)

Panel A. Revenues

Panel B. Expenditures



Notes: The Figure reports the point estimates and the respective 95% confidence intervals of OLS and instrumental variable (2SLS) coefficients obtained by regressing the change in natives' individual contributions on the change in immigrant inflows between 2013–2018 (Δm_{it}). We also report Anderson–Rubin robust confidence sets for 2SLS estimates using the geographic distribution of immigrants in 2005. Results are computed by trimming the distribution of contributions each year above the 99th percentile. All regressions include year dummies and interactions of MSA-level controls with year dummies. MSA controls include the share of local revenues that accrue to transfers from the central government, the share of working age population in the city, and the share of college workers. Estimates are weighted by the working-age population in 2013.

1.6 Mechanisms

So far, our results suggest that immigration has no effect on natives' (net) fiscal contributions. However, an important question to ask is if these findings are a result of small aggregate equilibrium effects (*e.g.*, labor market adjustments) or are a result of countervailing forces where indirect effects cancel out.

In this section, we address this question and explore the three mechanisms outlined in Section 2. We start by discussing the extent to which the fiscal effects presented before are mediated by changes in labor market outcomes for natives or changes in the price of non-tradable goods, particularly changes in property values and rents. We then discuss changes in the cost of provision and use of government services and the overall allocation of expenditures.

1.6.1 Labor Market Effects

As part of labor market equilibrium effects, we explore three possible mechanisms: changes in aggregate labor market outcomes (extensive margin), changes along the distribution of wages, and changes in the number of hours of work (intensive margin) and the quality of employment, measured by changes in the probability of being employed in the informal sector.

1.6.1.1 Labor Displacement and Average Wages

We start by testing if the increase in immigrant inflows generated labor displacement among native workers or affected average wages. This could explain our fiscal results, especially the negative effect on indirect taxes. We follow the ‘pure spatial approach’ literature and estimate the effect of immigration on natives’ labor market outcomes.²⁰ The baseline model regresses labor market outcomes (the probability of being employed, the probability of being unemployed, the probability of being active in the labor force, and log hourly wages) for natives on the fraction of immigrants in the working-age population of 2013, leveraging the pooled sample. We run a similar specification as in Eq. (8) but without including interactions of area-level characteristics in 2013 with time dummies.

Figure 1.4 presents OLS and 2SLS estimates for the impact of Venezuelan immigration on natives’ labor market outcomes. 2SLS estimates are presented for both the past settlements instrument (IV1) following Eq. (7) and the distance instrument (IV2) as in Eq. (9).²¹ Our specification does not suffer from a weak instrument problem since our instruments can predict immigration shares effectively. The Kleibergen-

²⁰ Our approach is similar to those implemented by Card (2007), Dustmann *et al.* (2013), and Peri & Yasenov (2019). In contrast to those papers which aggregate the outcomes of interest at the region by time level and regress differences over time in those outcomes in a particular skill group on changes in the fraction of immigrants, we exploit individual-level data.

²¹ Since we don’t estimate regressions using first differences, in Eq. (7) we use as “shift” the total stock of Venezuelan immigrants in Colombia in each year.

Paap F -stat for weak instruments ranges from 22.2 to 47, with a higher predictive capacity for the distance instrument.

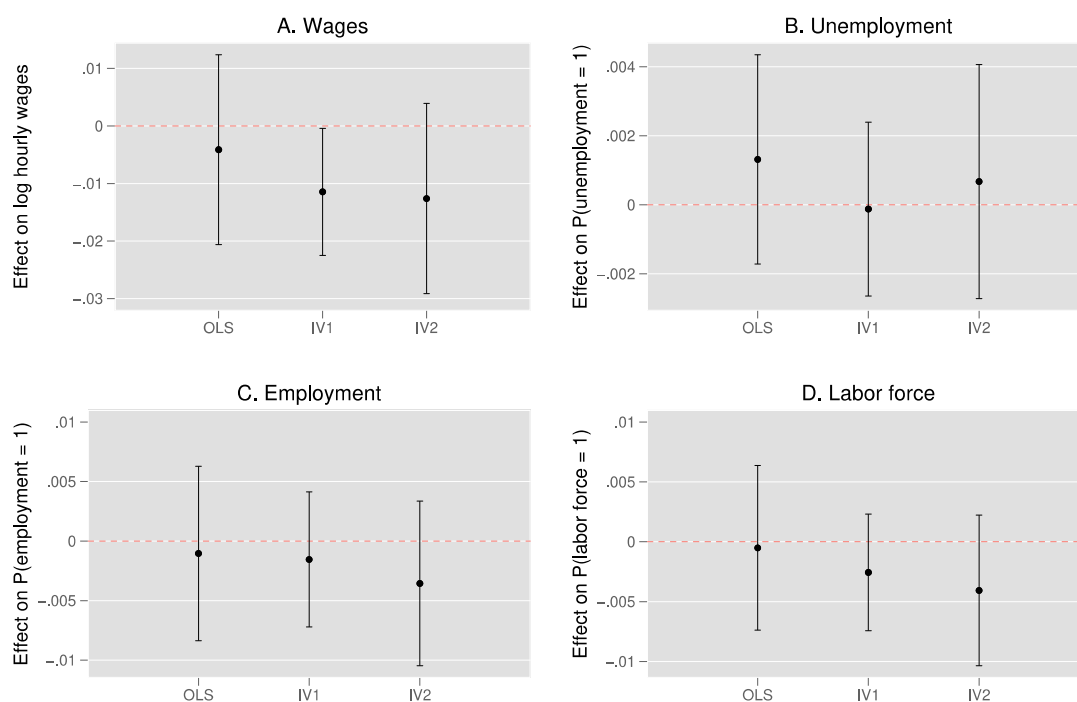
The OLS estimates in Figure 1.4 suggest that immigration from Venezuela is negatively correlated with natives' hourly wage and the probability of being employed and participating in the labor force, and positively correlated with the probability of being unemployed. However, these effects are statistically not distinct from zero. After accounting for the endogenous sorting of immigrants into areas and potential measurement error in the shares by instrumenting the fraction of immigrants with the past settlement and the distance-based instruments, the results show a negative and significant effect on wages only using the past settlement instrument (Panel A, IV1 estimate). On average, a 1 percentage point increase in the share of immigrants in the working-age population (for instance, going from 1% to 2%) is associated with a reduction in native wages of 1.1%. Placing the wage response into perspective, since the real hourly wage growth for working-age native workers between 2015 and 2018 (the high immigration period) was about 1.4% per year and the fraction of immigrants increased by 1 percentage point on average, the negative wage effect does not necessarily imply a decline in natives' real wages.

On the other hand, 2SLS estimates suggest that immigration from Venezuela does not seem to have influenced natives' aggregate

employment and labor participation outcomes between 2013 and 2018 (see IV1 and IV2 estimates on panels B, C, and D). These findings are robust to alternative specifications. We get comparatively similar results when we do not use sample weights, when controlling for a proxy of local economic activity (state GDP), and when building confidence intervals that are robust to weak instruments á la Anderson-Rubin as suggested by Andrews *et al.* (2019) (results available upon request).²²

Figure 1.4 Effect of immigration on main labor market outcomes

²² Our results are not directly comparable with some of the recent evidence of the effect of Venezuelan immigration on natives' labor market outcomes in Colombia as studies tend to differ in the definition of the fraction of immigrants, the period of analysis, the definition of Venezuelan immigrants, and the unit of analysis. Just as a reference, Carusso *et al.*, (2019) found large negative effects on wages and on the probability of being employed, and positive effects on the probability of being unemployed. Lebow (2021), using aggregate specifications at the MSA level between 2014-2019, only found negative effects on wages, and no effects on employment and unemployment. Delgado-Prieto (2021), looking at changes at the department-level, found negative effects on wages and employment in 2018 relative to 2015.



Notes: The Figure reports OLS and 2SLS coefficients and the respective 95% confidence intervals obtained by regressing natives' log wages (Panel A), unemployment (Panel B), employment (Panel C), and labor force participation status (Panel D) on the fraction of immigrants (\tilde{m}_{jt}). IV1 instruments \tilde{m}_{jt} with the past settlement instrument following Eq. (7). IV2 instruments \tilde{m}_{jt} with the distance instrument as defined in Eq. (9). All regressions include year and area fixed effects, individual controls (sex, age, age squared), and dummies for education achievement (less than high school, high school, some college, college graduates, and graduate degrees). Wages are computed for wage and salary workers and include the labor income of self-employed workers. The sample is restricted to natives aged 15 to 64 living in metropolitan areas (MSAs) from 2013-2018. To alleviate the potential impact of outliers, wages were computed by trimming the wage distribution by year at 0.5% and 99.5%. Wages are expressed in 2018 equivalent COP. Estimates are weighted by sampling weights. Standard errors are clustered at the metropolitan area level. The Kleibergen-Paap F -statistic ranges between 22.2 and 47.5 depending on the outcome variable, with higher predictive capacity for the distance instrument.

In Figures 1.12 and 1.13 in the Appendix, we explore whether labor market effects differ by workers' sex and skill, respectively. Results suggest that the impact on wages reported in Panel A, Figure 1.4 is mainly driven by the reduction in wages of low-skill males. For instance, the first panel in Figure 1.13 shows a meaningful negative and statistically

significant effect on the wages of low-skill workers (those with less than a high school diploma) using both the past settlement and the distance-based instruments. A 1 percentage point increase in the share of immigrants in the working-age population is associated with a reduction in real wages for low-skill natives between 1.8% and 2%, depending on the instrument used. The point estimates are in line with previous work from Lebow (2021) and Delgado-Prieto (2021), who estimate low-skill native responses to immigration from Venezuela between 1.4% and 2%.

Similarly, the first panel in Figure 1.12 shows a negative and significant effect on native males using the past settlement instrument. The coefficient implies a negative impact of immigration close to 2% for each 1 percentage point increase in immigration across metropolitan areas. Figures 1.12 and 1.13 also show no heterogeneity in the effects of immigration on natives' probabilities of employment, unemployment, and labor force participation by sex and skill. The effects of immigration seem to be concentrated on the wages of the low-skill males.²³

In Figures 1.14 and 1.15 in the Appendix, we further explore the effect on wages by economic sector and type of job (wage and salary

²³ We also replicate Figure A6 for females. Results for females confirm that only wages of low-skill male workers are being affected by the immigration supply shock. Our results show no effects of the Venezuelan immigration on the labor force participation of skilled native women with kids five years old or younger in the household between 2013 and 2018. This contrasts with previous findings in the literature studying the effect of low-skilled immigration in lowering the price of services such as housekeeping and the implications for labor force participation of high-skilled women (Cortes, 2008; Cortes & Tessada, 2011; Pedrazzi & Peñaloza-Pacheco, 2023).

workers, self-employed, government employees, and domestic workers), respectively. Figure 1.15 shows no wage effect for wage and salary workers (Panel A) but a negative and statistically significant effect on self-employed workers (Panel B). The effect on the wages of those self-employed could be explained by the fact that 59% are also low educated, and almost 90% are in the informal sector.²⁴ As a robustness check, we find, as expected, no effect on the wage of workers employed by the government. However, we do not see a decline in the wages of domestic workers. Because many immigrants are employed as domestic workers (see Lebow, 2021), we may have expected to find an effect for this group.

Taken together, the negative effect on natives' wages with relative effects concentrated in the wages of low-skill workers (mainly male) in the informal sector, in addition to no differential effect on employment and labor participation by skill, is consistent with a story of lack of downward wage rigidities, natives' inelastic labor supply, and imperfect substitution between immigrants and natives.

When natives' labor supply is inelastic, due to low reservation wages or to the fact that our period of the analysis is probably capturing the short-term effects, and there is no minimum wage protecting native workers in the informal sector, most of the adjustment to the immigrant

²⁴ Informality is defined as those in jobs with limited to no access to social security (pension or health benefits).

supply shock will be through wages. The degree to which firms' elasticity of labor demand will respond to wage changes in the informal sector will depend upon the degree of substitutability between natives and immigrants. Our findings of zero employment effects in response to changes in wages in the informal sector are consistent with the idea that immigrants and natives are imperfect substitutes. Therefore, most of the adjustment to immigration occurs through wages. These wage effects explain almost entirely the small decline in indirect taxes presented in Section 5.3.

1.6.1.2 Effects Along the Wage Distribution

We move to study the effects of immigration from Venezuela across natives' wage distribution. The fact that immigrants from Venezuela downgrade upon arrival, as depicted in Figure 1.16 in the Appendix, would lead one to think that the wage effect found in the previous section might not be constant across the wage distribution and would be particularly burdensome in segments of the labor market where the density of Venezuelan immigrants is higher than that of native workers.

We start by estimating log wages net of individual characteristics. Using the entire sample between 2013-2018 at the individual level, we regress log wages on sex, age, age squared, and dummies for education achievement (less than high school, high school, some college, college

graduates, and graduate degrees). We then get residual log hourly wages for different percentiles (5, 10, 15, ..., 95) for each MSA and year (\tilde{w}_{pjt}) and estimate:²⁵

$$\tilde{w}_{pjt} = \beta_p \tilde{m}_{jt} + \Omega_j + \Omega_t + \varepsilon_{pjt}, \quad (10)$$

where Ω_t and Ω_j are year and area fixed effects, respectively; \tilde{m}_{jt} is defined as in Eq. (8) and it's instrumented with the past-settlement and distance-based instruments described before.²⁶

Our specification is robust to immigrants' downgrading because it does not allocate immigrants into skill groups.²⁷ Downgrading is indeed an important issue when estimating the effects of immigration on natives' wages. Even though Venezuelan immigrants are, on average, more educated than natives, a large fraction of them remained undocumented and thus unable to work in the formal sector until the end of 2018 when

²⁵ In a similar application to the UK context, Dustmann *et al.* (2013) regress differences over time in percentiles of log wages across different regions on changes in the fraction of immigrants to natives, year dummies, and changes in the average age of immigrant and native workers in the regions, and the ratio of high (or intermediate) to low-educated native workers. Our empirical strategy differs from Dustmann *et al.* in two ways. First, we exploit individual-level data and use residual wages net of individual characteristics. Second, instead of using a first-difference estimator, we exploit variability across time within areas by using a two-way fixed effects estimator as defined in Eq. (10).

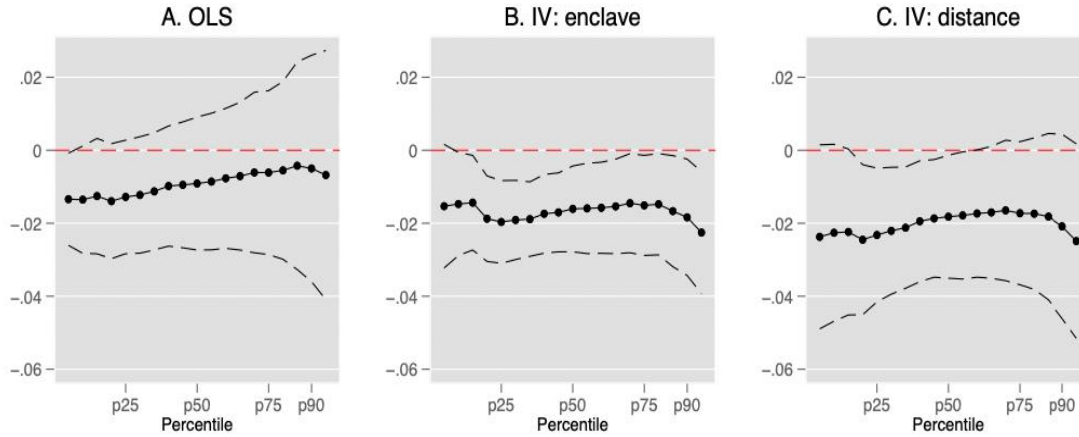
²⁶ Our approach, as in Dustmann *et al.* (2013), assumes the same effect of immigrants in any other segment in the wage distribution (or with different skills) on natives' wages. See Ottaviano & Peri (2012) for an estimate of the *total wage effect* of immigrants on natives' wages using a skill-cell approach. The skill-cell approach implies different effects depending not only on total immigration but also on the distribution of immigrants across skill groups. Such an approach, however, exploits variation in wages and immigrants across groups of workers with different skills at the national level, and not across local labor markets.

²⁷ For an example of the skill cell approach see Borjas (2003). For research using a mix approach (a combination of skill cell and spatial variation) see Dustmann & Glitz (2015).

the Colombian government extended the enrollment for the special working permit—PEP.

Figure 1.5 presents OLS and 2SLS estimates using the past settlement instrument (IV1) and distance-based instruments (IV2). The effects across the natives' wage distribution seem to mirror where immigrants from Venezuela are overrepresented. Figure 1.16 in the Appendix shows that Venezuelan immigrants are overrepresented below the 35th percentile in contrast to what one would have expected if they were compensated similarly to natives after accounting for immigrants' sex, age, and education (see green dashed line). Effects across the wage distribution using the distance-based instrument (Figure 1.5, Panel C) show a negative and significant effect on native wages between the 15th and 50th percentiles. The effects are larger at the lower part of the distribution. For instance, a 1 percentage point increase in the share of immigrants in the working-age population leads to a reduction of 2.4% and 1.9% in native hourly wages at the 20th and 50th percentile, respectively. Results are comparatively similar but slightly lower in magnitude using the past-settlement instrument (Figure 1.5, Panel B). In the case of the enclave instrument, there also appear to be effects in the upper part of the distribution. However, the effects above the 50th percentile are not distinct from zero at the 1% significance level.

Figure 1.5 Effect of Venezuelan immigration along natives' wage distribution



Notes: The Figure reports OLS and 2SLS coefficients and the respective 95% confidence intervals obtained by regressing residual log hourly wages for each percentile of the native wage distribution on the fraction of immigrants (\tilde{m}_{jt}) following Eq. (10). The enclave instrument is defined as in Eq. (7). The distance-based instrument is defined as in Eq. (9). Regressions are weighted by the number of observations in each quantile and area in 2013.

Does it make sense to find effects above the 50th percentile? As a reference, in our setting, the 50th percentile in 2018 is equivalent to the minimum wage, while the 75th percentile is equivalent to 1.5 times the minimum wage.²⁸ Finding some effects in the upper half of the wage distribution is not cause for concern for two reasons. First, while a requisite in the formal sector is to pay at least the minimum wage, the full cost of paying a minimum wage (including social security contributions, payroll taxes, and social benefits) amounts to twice the value of the minimum wage. Thus, it is unclear that wages at the 50th percentile

²⁸ As a comparison, Dustmann *et al.* (2013) find a negative impact of immigration below the 1st decile for the UK, which was at the same time around the minimum wage, but no significant effect up to about the 35th percentile. However, they find positive wage effects between the 40th and 90th percentile.

capture only workers in formal jobs. Second, there is long standing evidence that wages in the informal sector are indexed to the minimum wage and so it is used as reference by employers.

1.6.1.3 Intensive Margin and Labor quality

We explore two other possible sources of adjustment to increasing immigrant inflows from Venezuela: the intensive margin (by working fewer hours) or through changes in the quality of employment (changes in the probability of being employed in the informal sector). We estimate regressions using as dependent variables the hours worked by natives in the main occupation and different indicators of informality status.²⁹ Specifically, we use three related definitions of informality: 1) workers do not report pension contributions; 2) workers do not report health contributions; and 3) workers do not report both pension and health contributions.³⁰

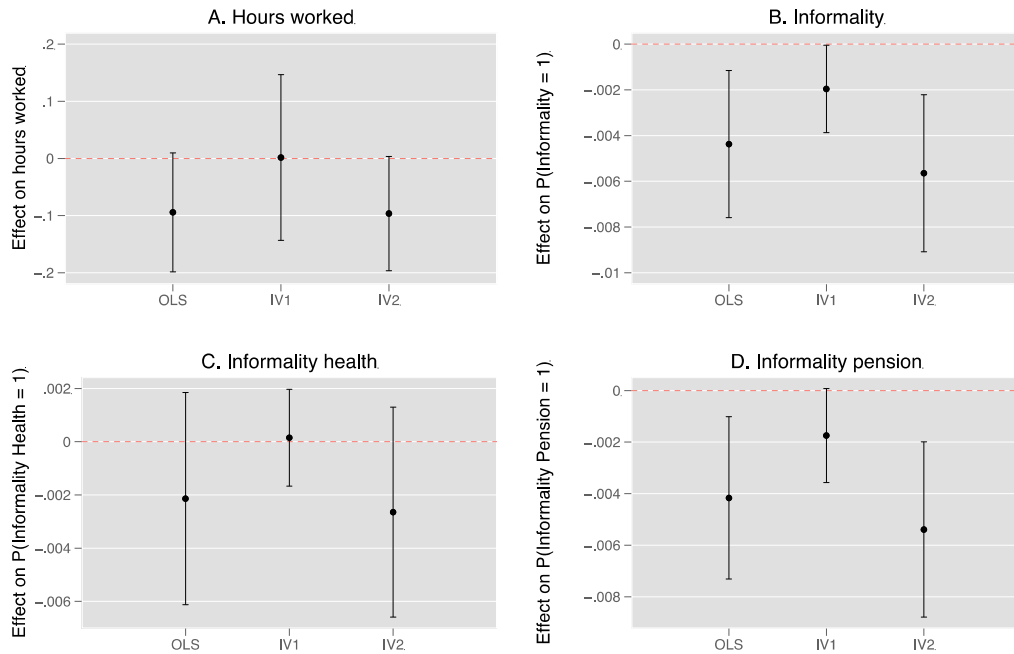
Figure 1.6 presents the results on the margins of adjustment. Panel A shows the effect on hours worked for native workers. We see a positive coefficient but not statistically distinct from zero, even after using both the past settlement and distance-based instrument. Regarding the probability of being employed informally, we see a small negative and significant effect

²⁹ We run a similar specification as in Eq. (8) but without including interactions of MSA-level characteristics in 2013 with time dummies.

³⁰ We classify as informal all those self-employed workers that don't make social security contributions and wage and salary workers for whom their employer is not making the legal contributions.

on informality (Panel B). A 1 percentage point increase in the fraction of immigrants reduces the probability of natives being employed in the informal sector between 0.2 and 0.5 percentage points, depending on the instrument used. These results do not necessarily imply a transition into the formal sector and out of informal jobs. However, it could suggest that natives are finding jobs where contributions to the pension system are more likely (Panel D). As most of the job opportunities for immigrants before the PEP was extended were in the informal sector, increasing competition might have induced some selection of natives into formal jobs. Indeed, among Venezuelan immigrants, 61% were working in the informal sector in 2014, increasing to 77% by 2017. These results add to our evidence on the effect on average wages for self-employed workers presented in Section 6.1.2.

Figure 1.6 Effect of immigration on the intensive margin and the quality of jobs



Notes: The Figure reports OLS and 2SLS coefficients and the respective 95% confidence intervals obtained by regressing natives' hours worked in the previous week in the main occupation and three measures of informality status on the fraction of immigrants (\tilde{m}_{jt}). IV1 instruments \tilde{m}_{jt} with the past settlement instrument following Eq. (7). IV2 instruments \tilde{m}_{jt} with the distance instrument as defined in Eq. (9). Informality status in Panel B refers to workers who do not report pension and health contributions. In panels C and D, informal is defined as not contributing to the health or pension system, respectively. All regressions include year and area fixed effects and individual controls (sex, age, age squared, and education). The sample is restricted to natives aged 15 to 64 living in metropolitan areas (MSAs) for 2013-2018. Estimates are weighted by sampling weights. Standard errors are clustered at the MSA-level. The F -statistic is 21.1 and 35.3 for the past settlement and distance instruments, respectively. The sample size is 1,328,858.

1.6.2 Property Values and Rents

Table 1.6 examines the effect of increasing immigrant inflows on property values [row (i)] and rents [row (ii)]. We look at variation across cities (Panel A) and neighborhoods within cities (Panel B). We use the geographic equivalent of a census tract (*sector urbano*) as our definition of neighborhood. Although only property values are used to determine the

property tax, the change in rents provides an additional measure of the market value of housing. Based on findings in the literature studying the Age of Mass Migration (Tabellini, 2020) and the Great Migration (Tabellini, 2019), it is not clear that one would find an effect on property values and rents, or that the effect goes in a particular direction, even in the short run. While large inflows of immigrants may increase the demand for housing, potentially increasing prices, at the same time, it could reduce the demand for housing among natives who, unwilling to live in larger immigrant enclaves, chose not to move to the city or move out of neighborhoods.

We regress $\ln(PRICE)_{kt}$ on the fraction of immigrants (\tilde{m}_{kt}), where $PRICE$ indicates the mean real value (rent) of owner-occupied (rental) units in MSA or neighborhood k in year t , net of observed housing quality characteristics.³¹ As before, we instrument the fraction of immigrants across cities using our enclave instrument. Now, because immigrants sort into neighborhoods, we also need to account for the endogenous allocation of immigrants within the city. We instrument the immigrant share in the tract using a modified version of our previous instrument. We interact

³¹ We start by regressing for each year log housing values on a set of housing quality characteristics (type of unit, number of rooms, structure and flooring material, utilities) and use the residual value as the dependent variable in our main regressions. These hedonic-type regressions account for changes in the quality-mix of properties over time. Since a household can rent either a whole unit or rooms within a unit, our rent price is defined as the average rent per room in the unit. So, the same unit can be owner-occupied and rent rooms. One drawback of the data is that housing values are based on owner self-reports.

predicted inflows of immigrants at the metro level for each year with the share of rooms in each tract in 2013.³² The idea behind is that tracts with a higher number of rooms will eventually accommodate a larger share of immigrants.

Focusing on the 2SLS results reported in Panel A, it seems that immigration had a negative and statistically significant impact on both property values and rents. A 1 percentage point increase in the share of immigrants in the area—relative to the local population in 2013—reduces housing prices and rents by 2.3% and 2.9%, respectively. When considering variation across neighborhoods (column 3, Panel B), these negative effects are more pronounced, reducing property values by 4.7% and rents by 3.9%. One crucial piece of information to keep in mind is that housing prices and rents were growing on average by 6.5% and 2.4% in real terms, respectively. At the same time, the fraction of immigrants increased by 0.66 percentage points on average between 2013 and 2018. Therefore, the negative effect does not necessarily imply a decline in prices and rents across areas. A more compelling interpretation of our results is that, at least in the case of property values, in the absence of immigration, areas that faced higher inflows would have experienced higher price growth. Despite the positive sign, this slower growth might explain why

³² Because the neighborhood identification in the GEIH is an orthogonal construction of the real, we are not able to use Census information for previous years to match to our data.

immigration did not significantly affect property tax revenues (Section 5.3).

Now, when we control for overall trends in the area, the negative neighborhood effect disappears (column 4, Panel B). By conditioning on tract and area by year fixed effects, the variation used for identification comes from changes in the fraction of immigrants in the tract over time, as compared to other tracts in the same area in a given year. In contrast to the results in Tabellini (2020), who finds that immigration from the Age of Mass Migration did not have a significant effect on property values in the US, some of our findings in the context of Venezuelan immigration suggest otherwise.³³

The negative effect on rents and property values may seem counterintuitive given that immigration leads to an increase in population and, therefore, demand for housing. Therefore, we test whether natives move away from immigrant-dense areas or neighborhoods. Natives may relocate across or within areas as a response to increasing immigrant inflows. For instance, if immigration affects natives' perceived value of local amenities, then immigrant inflows could induce outflows of natives

³³ The literature looking at the effects of black in-migration and racial desegregation has found also negative effects on property values and rents. For example, Boustan (2012) finds a decline in urban house prices and rents by 6% following the desegregation of public schools in the 1970s. Similarly, Tabellini (2019) finds that black in-migration to the North of the US between 1910 and 1930 reduced per capita property values by roughly 10%.

and lower housing prices and rents. Descriptive evidence shows that immigrants face multiple barriers when trying to find housing, from not having a credit history to being discriminated. This has resulted in about half of all immigrant households living in critical overcrowding, with most of them living in informal settlements.³⁴

Table 1.6 Effect of immigration on property values and rents

	Panel A. Variation across metropolitan areas			
	OLS		2SLS	
<i>(i) Property values</i>				
Immigrant share (\tilde{m}_{jt})	0.015 (0.014)		-0.023* (0.012)	
Kleibergen-Paap <i>F</i> -stat			18.327	
<i>(ii) Rents</i>				
Immigrant share (\tilde{m}_{jt})	-0.004 (0.008)		-0.029*** (0.039)	
Kleibergen-Paap <i>F</i> -stat			19.999	
Observations	138		138	
	Panel B. Variation across neighborhoods (<i>tracts</i>)			
	OLS	OLS	2SLS	2SLS
<i>(i) Property values</i>				
Immigrant share (\tilde{m}_{njt})	0.002 (0.005)	0.001 (0.003)	-0.047* (0.026)	0.056 (0.035)
Kleibergen-Paap <i>F</i> -stat			28.535	29.568
Observations	8,108	8,108	8,108	8,108
<i>(ii) Rents</i>				
Immigrant share (\tilde{m}_{njt})	-0.002 (0.000)	0.000 (0.001)	-0.039*** (0.011)	0.012 (0.011)
Area by Year FE		✓		✓
Kleibergen-Paap <i>F</i> -stat			53.822	31.080
Observations	8,178	8,178	8,178	8,178

³⁴ See *Encuesta de calidad de vida e integración de los migrantes venezolanos*: <https://migravenezuela.com/web/articulo/encuesta-de-calidad-de-vida-e-integracion-de-los-migrantes-venezolanos-en-colombia/2563>

Notes: The table reports the coefficients obtained by regressing the log of mean real property values [row (i)] and rents [row (ii)] on the fraction of immigrants (\tilde{m}_{kt}) between 2013–2018. Regressions in columns 1 and 3 include year and area (tract) fixed effects. Regressions in columns 2 and 4 control for tract and area by year fixed effects. The fraction of immigrants (\tilde{m}_{njt}) in Panel B is constructed by averaging immigrants in the tract using 2-year moving averages. Values are deflated using the price index for used housing from Colombia’s Central Bank and the rent index from the CPI reported by DANE. Results are computed by trimming the distribution of property values and rents each year at 0.5% and 99.5%. In all regressions, observations are weighted by the number of owner-occupied (or rental) units in the area or tract in 2013. Columns 1 and 3 report in parentheses robust standard errors. Columns 2 and 4 report robust standard errors clustered at the area level. *** Denotes significance at 1%, ** significance at 5% and * significance at 10%.

Table 1.8 in the Appendix tests for ‘native flight’. We find that changes in the fraction of immigrants within an area (neighborhood) lower the total native population in the area (neighborhood), consistent with natives moving away from immigrant-dense areas. These estimates are similar to those found by Accetturo *et al.* (2014) for a sample of Italian cities. As before, once we control for area-year-specific trends, we don’t see a difference between neighborhood and city average changes in the total native population.

1.6.3 Cost Effect

As discussed in Section 2, an increase in population size can affect the cost of government service provision. While we do not directly observe the unit cost of the multiple goods and services provided by local governments, we can use our estimates of the effect of immigrant inflows on the different expenditure groups to infer whether or not this mechanism is in place.

The fact that we do not see an increase in the per capita spending of public goods classified as non-rival in consumption or ‘pure’ (see Figure 1.3, Panel B) suggests no sign of costs being distributed among the larger population. On the contrary, spending classified as ‘pure’ public goods seems to share the characteristics of public goods that are rival in consumption and therefore tend to increase with population size. This may be the case of spending on general public administration and transfers of a general character between different levels of government.

In the case of health expenditures, as we show in Section 6.4., immigration did not significantly change the allocation of expenditures. Thus, part of the increase in natives’ health benefits may reflect an increase in the cost of providing the service. Changes in the cost could be driven by changes in the epidemiological profile of the population, explained by increasing immigrant inflows. For example, immigration could have impacted the cost of providing health services through changes in the population composition. Public records show that a significant fraction of Venezuelan immigrants are minors and over two-thirds of immigrant women arriving in Colombia are at their peak reproductive years. For instance, in 2013, there were only 25,500 Venezuelan-born immigrants under 18 years old, while the number of Venezuelan minors went to 507,679 in 2018, with 46% arriving in Colombia within the last 12 months. This, combined with limited information about the health records

of immigrants before arriving in the country, could have increased the health cost overall.

1.6.4 Composition and Service Use Effect

We now turn our attention to analyzing the effect that the increase in immigrant inflows has on the allocation of local government spending and service use. If we follow the findings in the literature, one may expect a negative effect for those expenditure items where interactions between groups are more salient or that imply larger transfers, such as productive public goods, education, or social protection. However, this is not necessarily true in our setting for at least three reasons. First, many local governments have been lenient—at least in their discourse—in helping in the humanitarian assistance and assimilation of immigrants. Second, local authorities tend to face less pushback from residents considering that Venezuelan immigrants are culturally not that different from natives. Third, local governments face limitations on local taxation combined with low budget flexibility, limiting the discretion of local authorities to change the allocation of public expenditures in the short run.

In Figure 1.7, we ask whether the change in the stock of Venezuelan immigrants affected the composition of local spending and the likelihood that natives rely on government-funded services or social programs. Panel A presents the change in the share of local spending by group (*e.g.*, public

goods, health, education, social protection, etc.). We estimate a version of Eq. (5) for each type of expenditure, and instead of using variation in the share of working-age migrants, we use the total population. Because some expenditures executed at the local level tend to be approved at higher levels of government (*e.g.*, transfers from the National government to pay for education and health), limiting the autonomy of local governments, Panel B provides results excluding transfers from the national government.

The results in Figure 1.7 show no significant change in local spending that could be explained by increasing immigrant inflows. However, a couple of things can be highlighted from the sign of some of the coefficients. First, after excluding transfers from the National government, the large and positive sign in health expenditures by local governments allows us to think that the increasing pressure on the health system from immigration, as many public officials have pointed out,³⁵ has required the fiscal effort of municipalities. Second, the shift in signs in both ‘congested’ public goods (*e.g.*, police services, transportation, housing development) and compulsory education, between Panel A and B, shows the large weight that government transfers have on local budgets, especially for smaller cities. Third, most of the humanitarian assistance for the vulnerable

³⁵ According to the Office of the Comptroller General, the debt with public hospitals only for urgent care provided to Venezuelan citizens exceeded \$407 billion in the last three years.

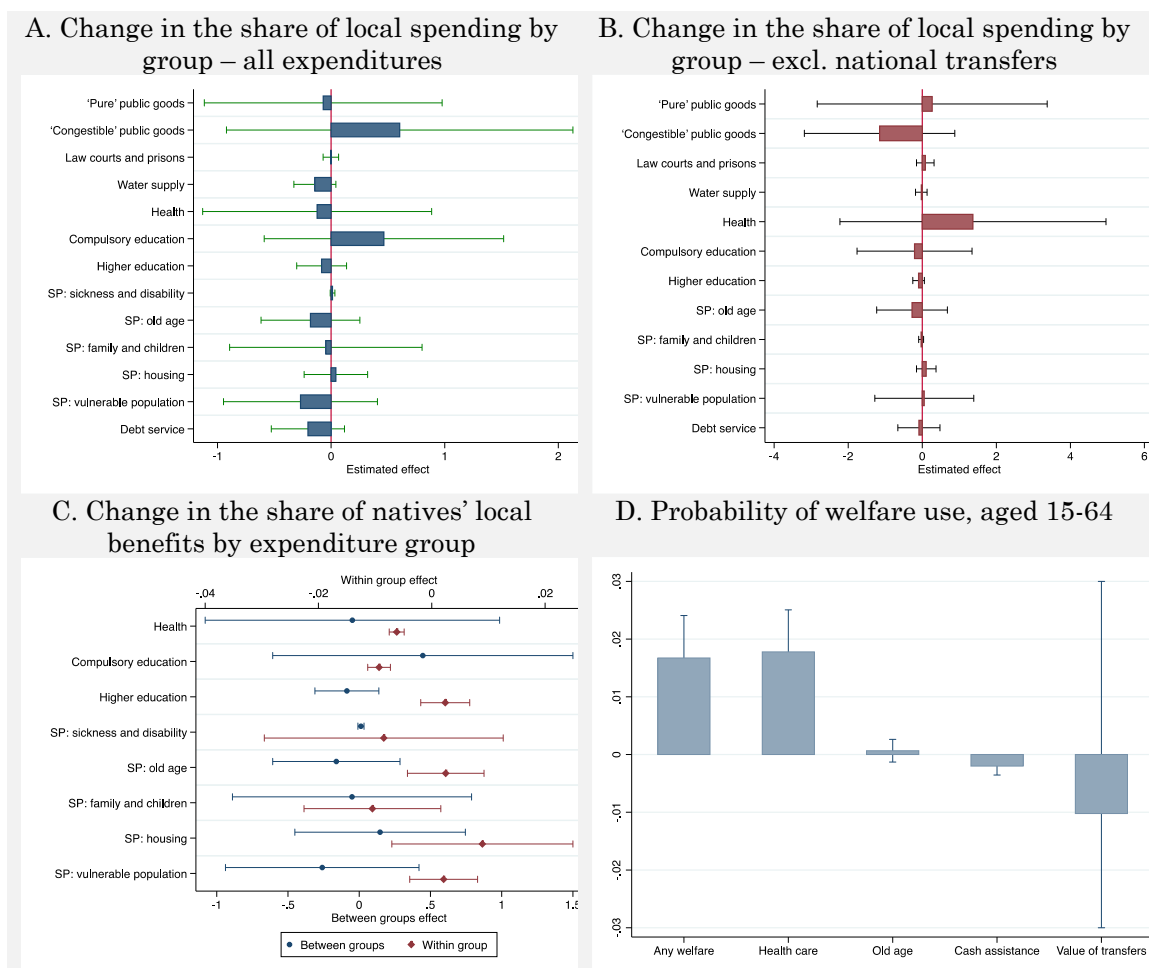
population, under which immigrants from Venezuela are beneficiaries, is being provided by the National government.

These results probably reflect low budget flexibility at the local level, which reduces the autonomy to reassign resources between expenditure groups. Therefore, in Panel C, we examine if, instead of shifts in overall spending, immigration might have affected the allocation of natives' local benefits. To that end, we exclude expenditures not derived from individual behavior and focus on health care, education, and social protection. We look at changes *between* expenditure groups and changes in spending *within* groups. Consistent with the results in Panel A and B, we do not see a significant change in the allocation of public spending across expenditure groups. However, when we look at the shares within groups, we do see a decline in natives' share in health and compulsory education. Comparing the *between* and *within* effects suggest that local governments have had to stretch their resources to meet the increasing demand. As Reina *et al.* (2018) pointed out, in the case of spending on compulsory education, the overall resources transferred to regional and local governments are allocated essentially based on the cost of the payroll, which initially reduces the need for additional spending.

Finally, in Panel D, we estimate whether natives are more or less likely to receive welfare due to increasing immigrant inflows. We examine

whether individuals claim welfare benefits, such as subsidized health care, conditional cash transfers or other cash assistance, and benefits for older adults. We also check if immigration changed the cash value of transfers received. We estimate a linear probability model and control for observable individual characteristics that may influence welfare take-up, as well as area and year fixed effects.

Figure 1.7 Composition and service use effect



Notes: The Figure reports the point estimates and the respective 95% confidence intervals of 2SLS coefficients for the fraction of immigrants (\tilde{m}_{jt}). The dependent variable in Panel A is the share of local public spending in each expenditure group. In Panel B, the shares exclude transfers from the national government. In Panel C, only information on spending allocated to natives is used to construct the *between* and *within* shares. All regressions include year dummies and interactions of

MSA-level controls at their 2013 level with year dummies. MSA controls include the share of local revenues that accrue to transfers from the central government, the share of the working-age population in the city, the share of college-educated workers, and the log of mean labor income. All estimates are weighted by the total population in the area in 2013. Panel D reports the probability of taking-up welfare, using as dependent variables dummies that indicate self-reported information on receiving any welfare, subsidized health, old age benefits, or cash assistance, and the real value of cash transfers. Regressions in Panel D control for year and area fixed effects, age, sex, family size, monthly wage, nonlabor and non-transfer related income.

Immigration appears to have increased welfare take-up by natives. For instance, a 1 percentage point increase in the fraction of working-age immigrants, relative to the local population in 2013, raises the probability of claiming welfare by 1.7 percentage points. This effect is driven by a higher probability of being enrolled in the subsidized health system. Although the results show a slight drop in the probability of receiving any type of cash assistance, the cash value does not seem to have changed. These findings are consistent with a decline in the probability of working informally; however, we believe these effects are too small to explain any effect on fiscal contributions.

The effects on subsidized health take-up are somewhat puzzling, especially since we do not find an effect on informality measured by contributions to the health system. We provide one interpretation for these findings. Considering that between 2013 and 2018 the total health coverage went from 96% to 94.7%, and the subsidized regime did not grow in the number of affiliates, these results suggest that areas with higher inflows of immigrants experienced a slower decline—perhaps even a slight increase—in coverage. Overall, the increase in the probability of natives

relying upon subsidized health care is consistent with the increase in per capita health expenditures shown in Figure 1.3.

1.7 Discussion

What do these results imply for the effect of immigration on public finances? In this paper, we have tried to address a valid concern of traditional accounting estimates of the net fiscal effect of immigration, which estimate individual tax contributions and benefits, ignoring labor market equilibrium effects, changes in the price of non-tradable goods, and assuming no interdependency in the level of public goods that both natives and migrants consume. If these effects are present, not accounting for them in traditional fiscal estimates will most likely underestimate the fiscal effects if immigrants' tax contributions are not enough to cover the benefits they receive or overestimate their effect if they generate a fiscal surplus.

Overall, our reduced-form findings indicate that immigration did not significantly change natives' average net fiscal contributions between 2013 and 2018. This result is supported by our analysis of the mechanisms potentially affecting natives' contribution. Not only do we not find labor displacement effects, but there is also no evidence of changes in the allocation of public expenditures, a differential increase in housing values, or an increase in the cost of public services. Even when we show that

increasing immigrant inflows did induce a decline in the wages at the bottom of the wage distribution, particularly for relatively low-skilled workers, this group of workers is more likely not to meet the threshold required to pay income taxes or wealth tax. They are also more likely to be self-employed or work in the informal sector, meaning they usually contribute less to the health and pension systems. Because these add up to be a large share of overall contributions, results showing a slight decline in contributions to value-added and other indirect taxes due to lower wages at the bottom of the distribution do not significantly change natives' average contributions.

Estimates that adjust for aggregate general equilibrium effects that do not consider the heterogeneity in the labor market and the rate at which natives use government services may induce a more considerable bias in fiscal estimates. Even in contexts that have experienced a decline in employment, the job loss does not mean that natives would have paid more taxes and received fewer benefits in the absence of immigration. In addition, looking at changes in average wages does not truly capture the potential impact on fiscal estimates. Differential responses in the bottom and upper parts of the wage distribution might have very different effects on fiscal contributions.

It is essential to acknowledge that these effects are static and backward-looking in nature. Any intent to obtain a complete picture of the fiscal effects of immigration would require an evaluation—under a set of assumptions—about the future life trajectories of immigrants and their descendants. Our results have implications for forward-looking analysis. Particularly in settings where additional immigrant inflows are projected. Forward-looking estimates usually assume that individual tax contributions and benefits (be it natives or immigrants) will behave similarly (absent a policy change) to what an average individual with similar characteristics displays in the baseline year. Thus, understanding whether we need to account for indirect effects in the baseline calculations has implications for estimates of the lifetime net fiscal impacts.

There is also the political economy of immigrants' fiscal effects. Public opinion often tends to emphasize the negative indirect effects but overlook those positive impacts. Because estimates of the net fiscal impact of immigration are often used to inform immigration policy, it is important to have a complete understanding of direct and indirect effects.

Finally, we would like to acknowledge some limitations of our results. On the one hand, we do not directly account for changes in revenues from capital income caused by increasing immigrant inflows. The argument is that profit-maximizing firms hiring additional workers would

have had to employ additional capital, creating additional capital income. Omitting this effect has important implications for fiscal estimates and is a usual source of bias in cash flow accounting exercises, as shown by Clemens (2022). This would affect both natives' personal and corporate income tax contributions. On the other hand, our results assume no fiscal cross-subsidization across areas. If immigration, for example, leads to higher spending in one area that is financed with possibly higher taxes paid by other regions, then we might not be getting an accurate picture of the full effects of immigration. Another limitation is that we don't examine whether it is possible that immigration affected tax rates themselves through the political process. One promising area for future research could be to study and decompose the effect on total taxes paid into a tax rate effect and a tax base effect.

1.8 Concluding Remarks

While immigration in Colombia is estimated to generate a small fiscal benefit, consistently, more than half of residents consider immigration to be a burden for public finances. Even when we look only at Venezuelan immigrants, who have a negative fiscal position, their effect as a fraction of total economic activity is rather small. Now, a relatively common critique of these types of estimates that are produced using a static fiscal accounting analysis is that they usually miss general equilibrium effects

that are mediated through factor markets (labor and capital), changes in the cost of providing public services, and mean service used by both natives and immigrants.

As a result, the estimates produced by this partial-equilibrium approach might likely be underestimating the effect of Venezuelan immigration to Colombia between 2013 and 2018. Using variation in the inflows of immigrants received by Colombian metro areas during this period, we find no evidence that increasing immigrant inflows to Colombia lowered the average per capita contributions of natives. This is consistent with no significant decline in employment, average wages in the upper half of the wage distribution, or hours worked (as part of labor market adjustment), and no real decline in property values or change in the composition of local public spending. In addition, we find no evidence of costs being distributed among a larger population, suggesting that the composition of public expenditures is biased towards services that tend to increase with population size.

The findings in this paper may be specific to the Colombian context and may not hold necessarily after 2018. However, they may still be relevant for understanding the fiscal effects of immigration in developing countries and how local governments adjust to immigrant shocks in the short run. The results presented here also show the need to account for

second-order effects to provide a complete view of immigrants' fiscal effects.

REFERENCES

- Accetturo, A., Manaresi, F., Mocetti, S., & Olivieri, E. (2014). Don't Stand So Close to Me: The Urban Impact of Immigration. *Regional Science and Urban Economics*, 45, 45–56.
- Alesina, A., Baqir, R., & Easterly, W. (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics*, 114(4), 1243–1284.
- Altonji, J. G., & Card, D. (1991). The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives. In J. M. Abowd, & R. B. Freeman, *Immigration, Trade, and the Labor Market* (pp. 201–234). Chicago, IL: University of Chicago Press.
- Andrews, I., Stock, J. H., & Sun, L. (2019). Weak Instruments in IV Regression: Theory and Practice. *Annual Review of Economics*, 11(1), 727–753.
- Auerbach, A., & Oreopoulos, P. (1999). Analyzing the Fiscal Impact of US Immigration. *American Economic Review*, 89(2), 176–180.
- Aydemir, A., & Borjas, G. J. (2011). Attenuation Bias in Measuring the Wage Impact of Immigration. *Journal of Labor Economics*, 29(1), 69–113.
- Bayer, P., Ross, S., & Topa, G. (2008). Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116(6), 1150–1196.
- Blau, F. (1984). The Use of Transfer Payments by Immigrants. *Industrial and Labor Relations Review*, 37(2), 222–239.
- Bonilla-Mejía, L., Morales, L., Hermida-Giraldo, D., & Flórez, L. (2020). El Mercado Laboral de los Inmigrantes y no Inmigrantes. Evidencia de la Crisis Venezolana de Refugiados. Borradores de Economía 1119, Banco de la Republica de Colombia.
- Borjas, G. J. (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.
- Borjas, G. J., & Hilton, L. (1996). Immigration and the Welfare State: Immigrant Participation in Means-Tested Entitlement Programs. *Quarterly Journal of Economics*, 111(2), 575–604.
- Boustan, L. P. (2012). School Desegregation and Urban Change: Evidence from City Boundaries. *American Economic Journal: Applied Economics*, 4(1), 85–108.

- Calderón-Mejía, V., & Marinescu, I. (2012). The Impact of Colombia's Pension and Health Insurance Systems on Informality. IZA Discussion Paper No. 6439.
- Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1), 22–64.
- Card, D. (2007). How Immigration Affects U.S. Cities. *Working Paper, Centre for Research and Analysis of Migration*. London, UK.
- Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2), 1–21.
- Card, D., & Peri, G. (2016). Immigration Economics by George J. Borjas: A Review Essay. *Journal of Economic Literature*, 54(4), 1333–1349.
- Caruso, G., Gomez-Canon, C., & Mueller, V. (2019). Spillover Effects of the Venezuelan Crisis: Migration Impacts in Colombia. *Oxford Economic Papers*, 1–25.
- Clemens, M. A. (2022). The Fiscal Effect of Immigration: Reducing Bias in Influential Estimates. IZA Discussion Paper No. 15592.
- Cortes, P. (2008). The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data. *Journal of Political Economy*, 116(3), 381–422.
- Cortes, P., & Tessada, J. (2011). Low-Skilled Immigration and the Labor Supply of Highly Skilled Women. *American Economic Journal: Applied Economics*, 3(3), 88–123.
- Dahlberg, M., Edmark, K., & Lundqvist, H. (2012). Ethnic Diversity and Preferences for Redistribution. *Journal of Political Economy*, 120(1), 41–76.
- Delgado-Prieto, L. (2022). Immigration, Wages, and Employment under Informal Labor Markets. UC3M Working papers, Economics 35664, Universidad Carlos III de Madrid, <https://e-archivo.uc3m.es/bitstream/handle/10016/35664/we2212.pdf?sequence=1>.
- Dustmann, C., Frattini, T., & Preston, I. P. (2013). The Effect of Immigration along the Distribution of Wages. *The Review of Economic Studies*, 80(1), 145–173.
- Dustmann, C., & Frattini, T. (2014). The Fiscal Effects of Immigration to the UK. *Economic Journal*, 124(580), F593–F643.

- Dustmann, C., & Glitz, A. (2015). How Do Industries and Firms Respond to Changes in Local Labor Supply? *Journal of Labor Economics*, 33(3), 711–750.
- Dustmann, C., Schönberg, U., & Stuhler, J. (2016). The Impact of Immigration: Why Do Studies Reach Such Different Results? *Journal of Economic Perspectives*, 30(4), 31–56.
- Dustmann, C., Schönberg, U., & Stuhler, J. (2017). Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment. *The Quarterly Journal of Economics*, 132(1), 435–483.
- Foged, M., & Peri, G. (2016). Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data. *American Economic Journal: Applied Economics*, 8(2), 1–34.
- Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8), 2586–2624.
- Hansen, J., & Lofstrom, M. (2003). Immigrant Assimilation and Welfare Participation: Do Immigrants Assimilate into or out of Welfare? *Journal of Human Resources*, 38(1), 74–98.
- Heath, R. (2018). Why Do Firms Hire Using Referrals? Evidence from Bangladeshi Garment Factories. *Journal of Political Economy*, 126(4), 1691–1746.
- Jaeger, D., Ruist, J., & Stuhler, J. (2019). Shift-Share Instruments and Dynamic Adjustments: The Case of Immigration. NBER Working Paper 24285.
- Lebow, J. (2021, September). The Labor Market Effects of Venezuelan Migration to Colombia: Reconciling Conflicting Results. Working Paper, https://static1.squarespace.com/static/6064ff1093aaa5277d9d8df9/t/615c5d64041de46ded2ec9fd/1633443180371/lebow2021_labor_market_effects.pdf.
- Manski, C. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3), 531–542.
- Mesa-Guerra, C. A., & Ramírez-Tobón, T. (2022). Estimating the Effect of Immigration on Public Finances: Evidence from the Influx of Venezuelan Migrants to Colombia. Working Paper, Available at SSRN 3927396.

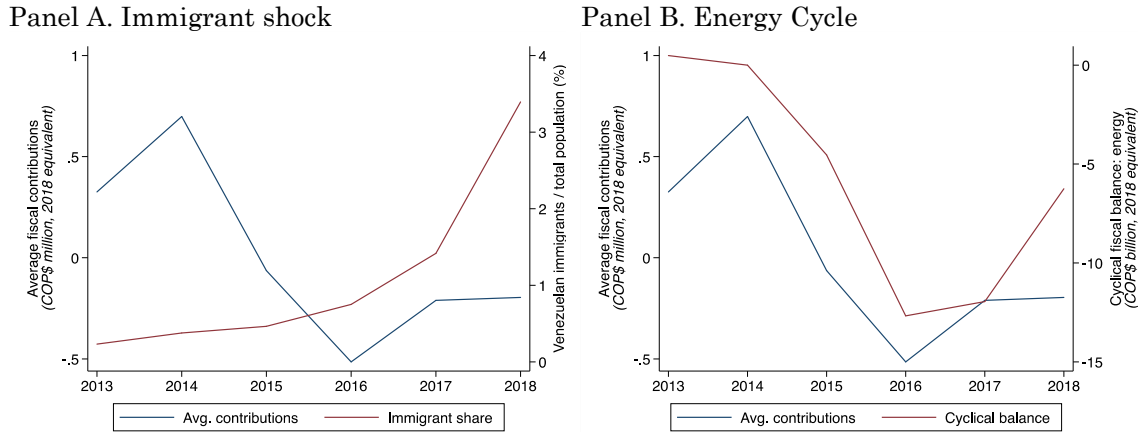
- Munshi, K., & Rosenzweig, M. (2018). Ethnic Diversity and the Under-Supply of Local Public Goods. Mimeo, University of Cambridge.
- National Academies of Sciences, E. a. (2017). *The Economic and Fiscal Consequences of Immigration*. Washington, DC: The National Academies Press.
- National Research Council. (1997). *The New Americans: Economic, Demographic, and Fiscal Effects of Immigration*. Washington, D.C.: National Academy Press.
- OECD. (2013). *International Migration Outlook 2013*. Paris: OECD publishing.
- OECD/ILO. (2018). *How Immigrants Contribute to Developing Countries' Economies*. Paris: OECD Publishing.
- Ottaviano, G. I., & Peri, G. (2012). Rethinking the Effect of Immigration on Wages. *Journal of the European Economic Association*, 10(1), 152–197.
- Pearl, J. (2013). Linear Models: A Useful “Microscope” for Causal Analysis. *Journal of Causal Inference*, 1(1), 155–170.
- Pedrazzi, J., & Peñaloza-Pacheco, L. (2023). Heterogeneous Effects of Forced Migration on the Female Labor Market: The Venezuelan Exodus in Colombia. *Journal of Development Studies*, 59(3), 324–341.
- Peri, G., & Yasenov, V. (2019). The Labor Market Effects of a Refugee Wave. *Journal of Human Resources*, 54(2), 267–309.
- Preston, I. (2014). The Effect of Immigration on Public Finances. *Economic Journal*, 124(580), F569–F592.
- Reina, M., Mesa, C. A., & Ramírez, T. (2018). Elementos para una Política Pública Frente a la Crisis de Venezuela. *Cuadernos de Fedesarrollo*, 69, p. 192.
- Steiner, R. & Cañas, A. (2014). Tributación y Equidad en Colombia. In Montenegro & Meléndez (ed.), *Equidad y Movilidad Social: Diagnósticos y Propuestas para la Transformación de la Sociedad Colombiana*, Universidad de los Andes-Departamento Nacional de Planeación, 415– 480.
- Tabellini, M. (2019). Racial Heterogeneity and Local Government Finances: Evidence from the Great Migration. Harvard Business School Working Paper, No. 19-006.

- Tabellini, M. (2020). Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration. *Review of Economic Studies*, 87(1), 454–486.
- Weber, S., & Péclat, M. (2017). A Simple Command to Calculate Travel Distance and Travel Time. *The Stata Journal*, 17(4), 962–971.

APPENDIX

A. Additional Figures and Tables

Figure 1.8 Fiscal contributions, immigrant shock, and energy cycle



Notes: The Figure compares the evolution between 2013 and 2018 of average net fiscal contributions for working-age natives to the immigrant share (Panel A) and the cyclical fiscal balance for the energy sector (Panel B). The data for the energy cyclical balance comes from public records published by the Ministry of Finance and measures the change in fiscal revenues caused by the difference between the observed price and the long-term price of crude oil of the previous period.

Figure 1.9 Directed Acyclic Graph (DAG) of the Effect of Immigration on Natives' Fiscal Contributions

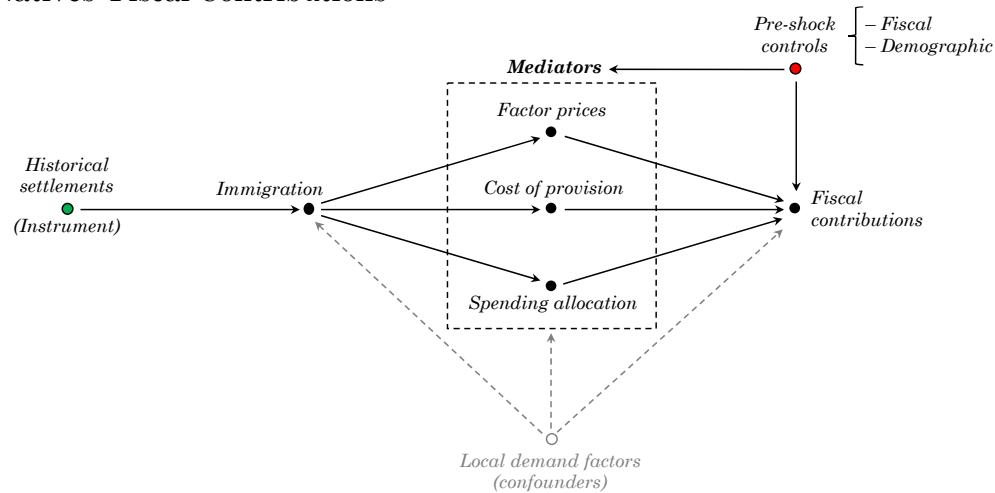
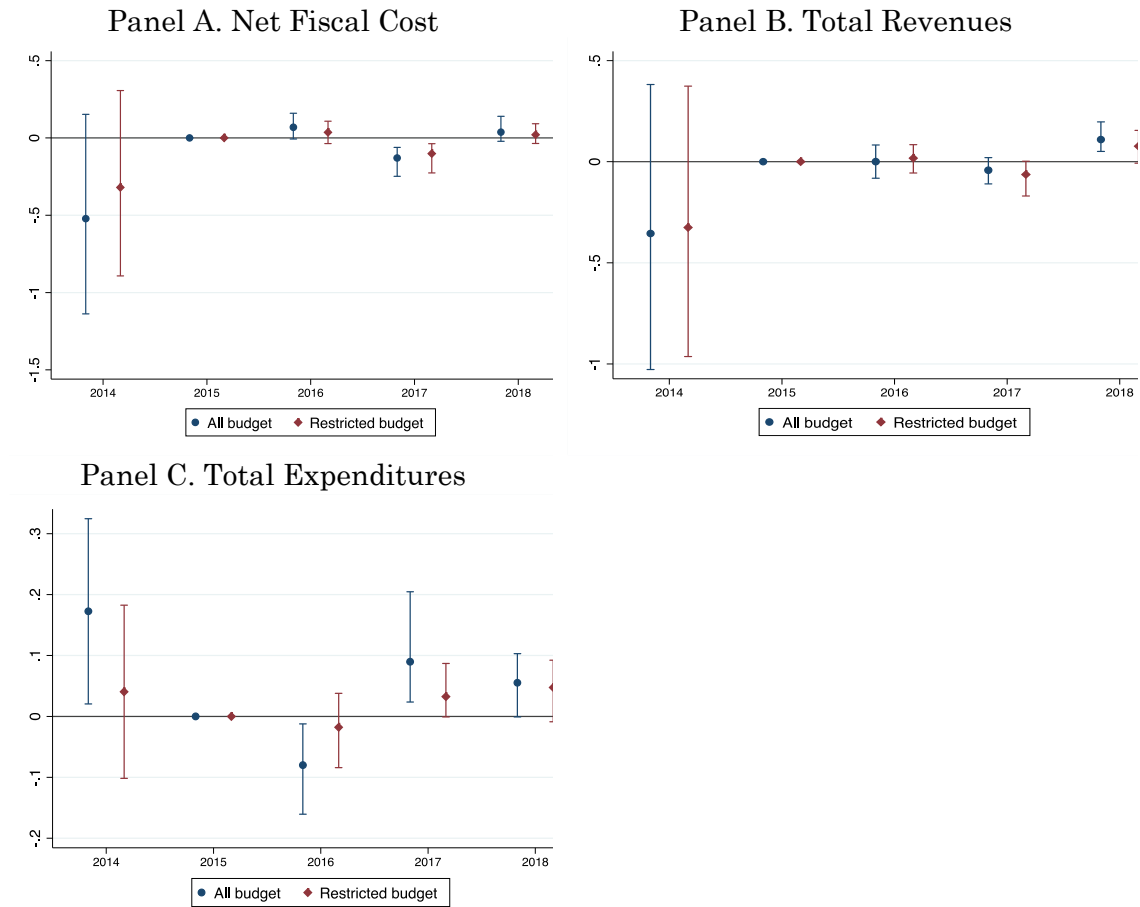
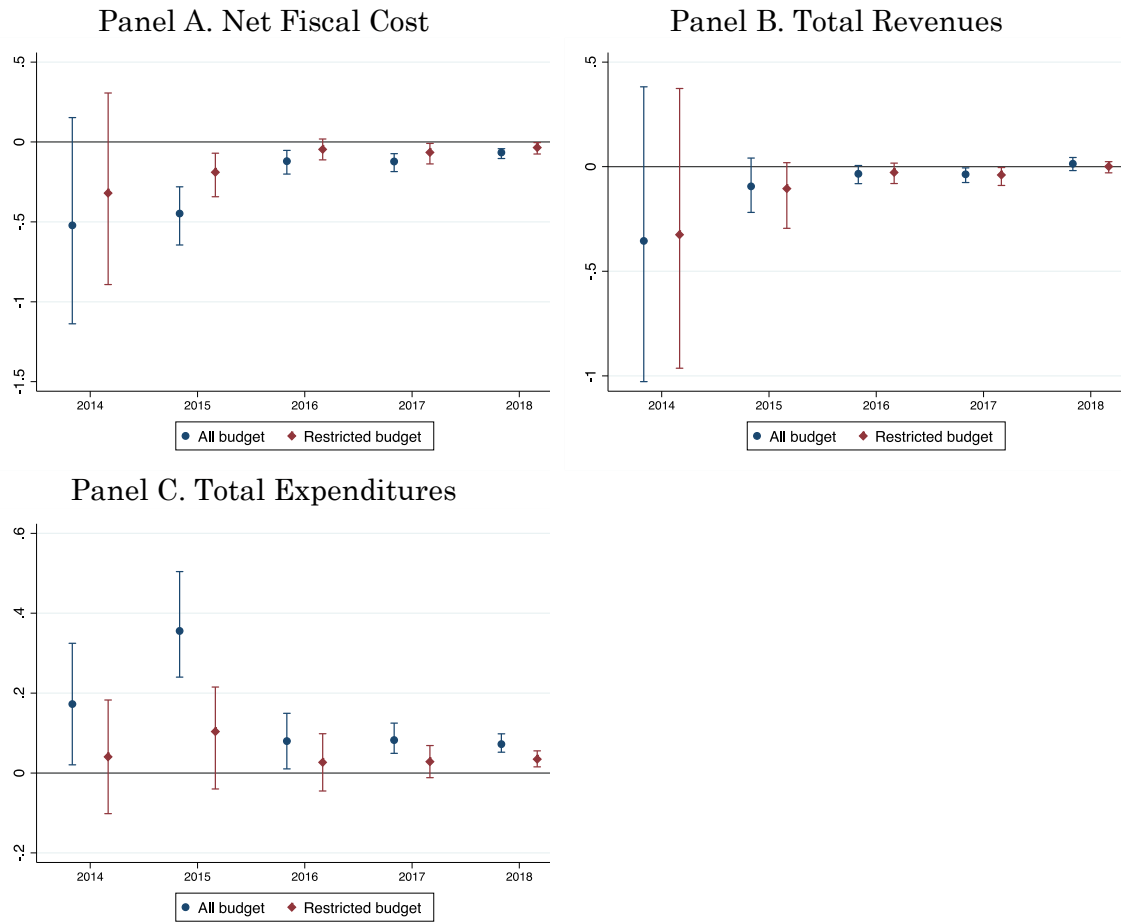


Figure 1.10 Year-by-year estimates of the effect of immigration on natives' fiscal contributions



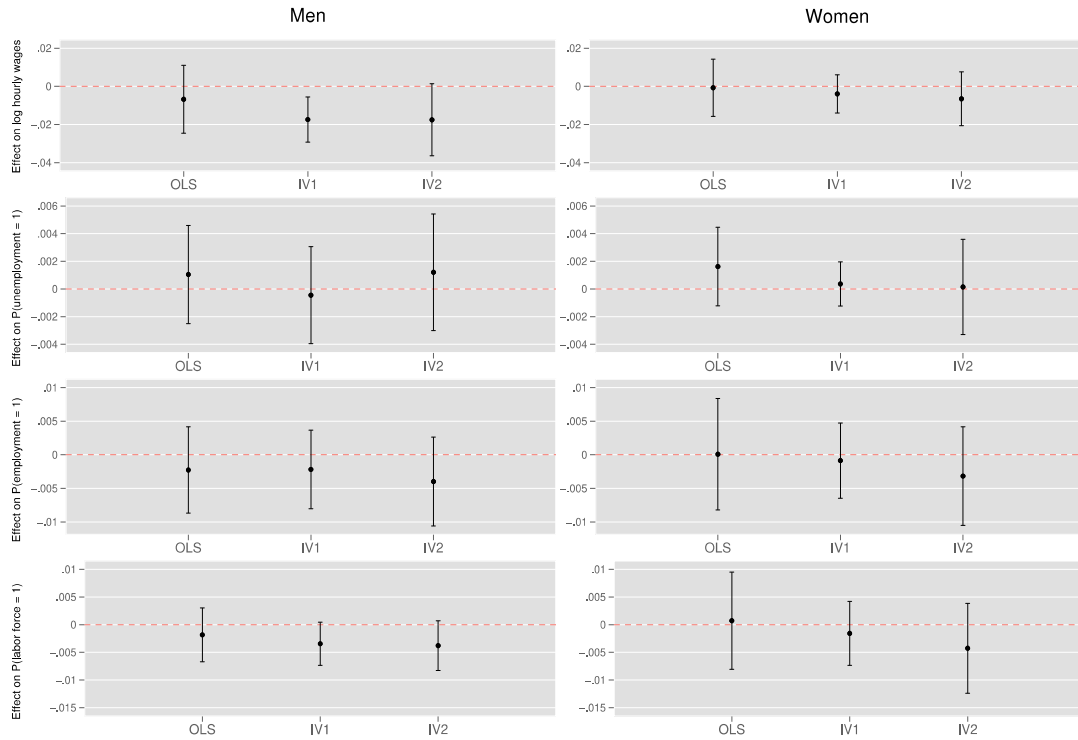
Notes: The Figure reports the year-by-year coefficients obtained by regressing the change in net fiscal contributions (panel A), revenues (Panel B), and expenditures (Panel C) on the change in immigrant inflows (Δm_{it}). All regressions control for the share of local revenues that accrue to transfers from the central government, the share of expenditures in public goods, the share of working age population in the city, the share of college workers, and the share of workers employed in manufacturing. Anderson–Rubin confidence sets are presented. Coefficients in 2015 are set to zero as a result of weak instruments. Estimates are weighted by the working-age population in 2013.

Figure 1.11 Estimates of the effect of immigration on natives' fiscal contributions relative to 2013



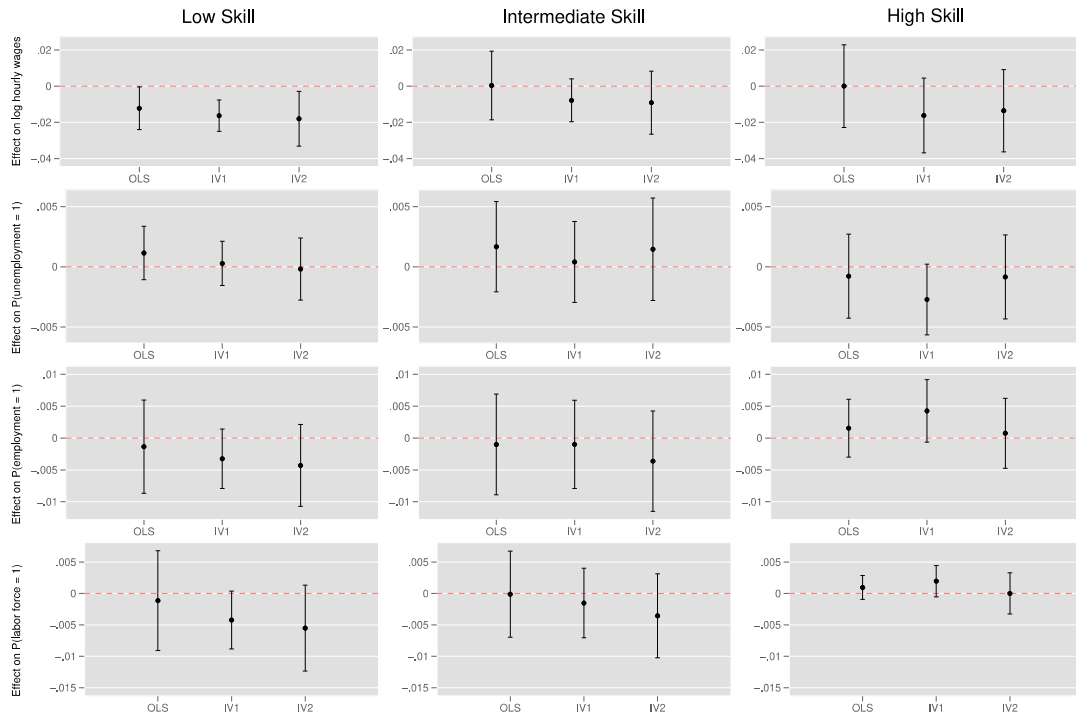
Notes: The Figure reports the coefficients obtained by regressing the change in net fiscal contributions (panel A), revenues (Panel B), and expenditures (Panel C) on the change in immigrant inflows relative to the local population in 2013. All regressions control for the share of local revenues that accrue to transfers from the central government, the share of expenditures in public goods, the share of working age population in the city, the share of college workers, and the share of workers employed in manufacturing. Anderson–Rubin confidence sets are presented. Estimates are weighted by the working-age population in 2013.

Figure 1.12 Effect of immigration on main labor market outcomes by sex



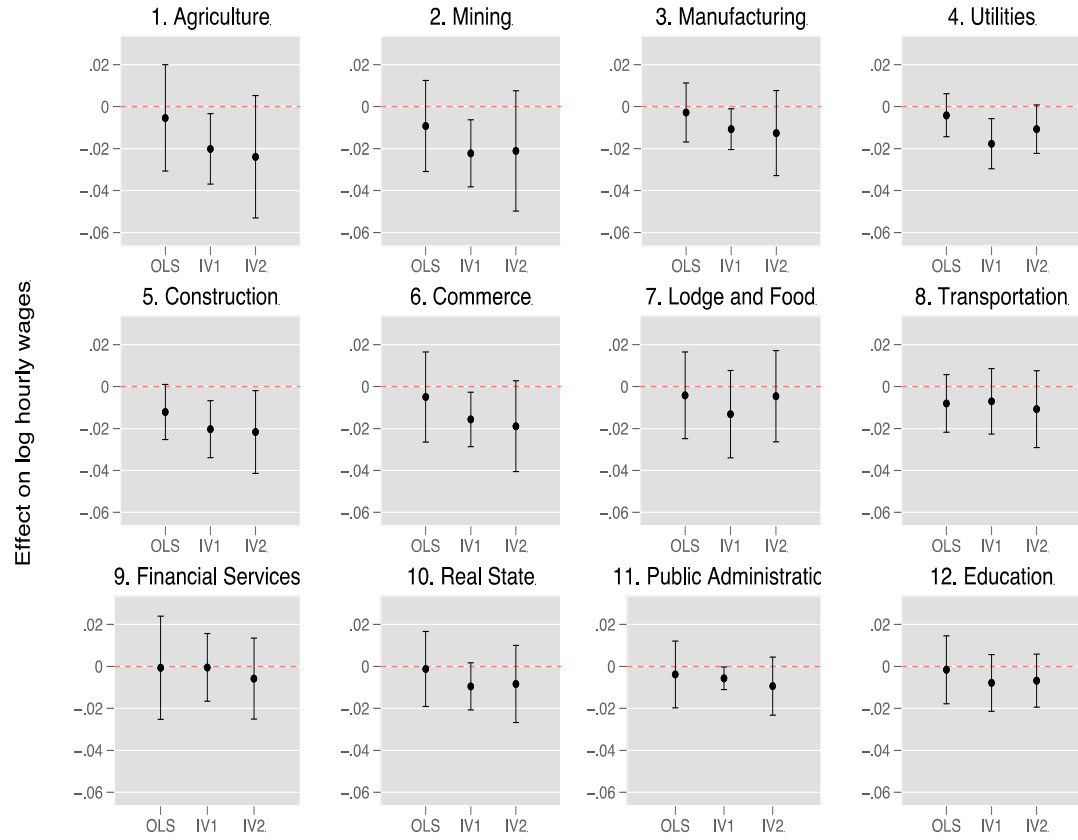
Notes: The Figure reports OLS and 2SLS coefficients and the respective 95% confidence intervals obtained by regressing natives' log wages, unemployment, employment, and labor force participation status on the fraction of immigrants (\tilde{m}_{jt}) separated by sex. IV1 instruments \tilde{m}_{jt} with the past settlement instrument as defined in Eq. (7). IV2 instruments \tilde{m}_{jt} with the distance instrument as defined in Eq. (9). All regressions include year and area fixed effects, individual controls (sex, age, age squared), and dummies for education achievement (less than high school, high school, some college, college graduates, and graduate degrees). Wages are computed for wage and salary workers and include the labor income of self-employed workers. The sample is restricted to natives aged 15 to 64 living in metropolitan areas (MSAs) from 2013-2018. To alleviate the potential impact of outliers, wages were computed by trimming the wage distribution by year at 0.5% and 99.5%. Wages are expressed in 2018 equivalent COP. Estimates are weighted by sampling weights. Standard errors are clustered at the metropolitan area level.

Figure 1.13 Effect of immigration on main labor market outcomes by skill



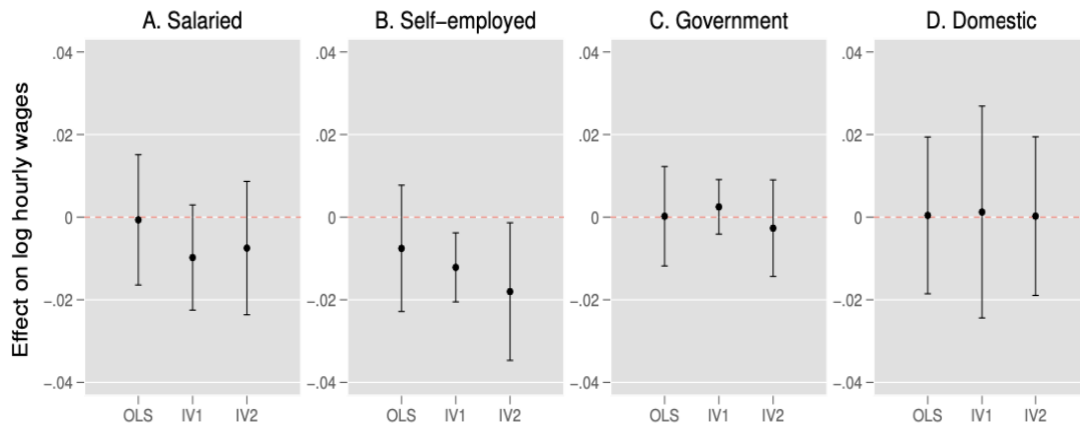
Notes: The Figure reports OLS and 2SLS coefficients and the respective 95% confidence intervals obtained by regressing natives' log wages, unemployment, employment, and labor force participation status on the fraction of immigrants (\tilde{m}_{jt}) separated by skill. IV1 instruments \tilde{m}_{jt} with the past settlement instrument as defined in Eq. (7). IV2 instruments \tilde{m}_{jt} with the distance instrument as defined in Eq. (9). Low skill: those with less than high-school degrees. Intermediate skill: those with high school degrees or technical degrees. High skill: those with a bachelor's or more. All regressions include year and area fixed effects, individual controls (sex, age, age squared), and dummies for education achievement (less than high school, high school, some college, college graduates, and graduate degrees). Wages are computed for wage and salary workers and include the labor income of self-employed workers. The sample is restricted to natives aged 15 to 64 living in metropolitan areas (MSAs) from 2013-2018. To alleviate the potential impact of outliers, wages were computed by trimming the wage distribution by year at 0.5% and 99.5%. Wages are expressed in 2018 equivalent COP. Estimates are weighted by sampling weights. Standard errors are clustered at the metropolitan area level.

Figure 1.14 Effect of immigration on wages by sector



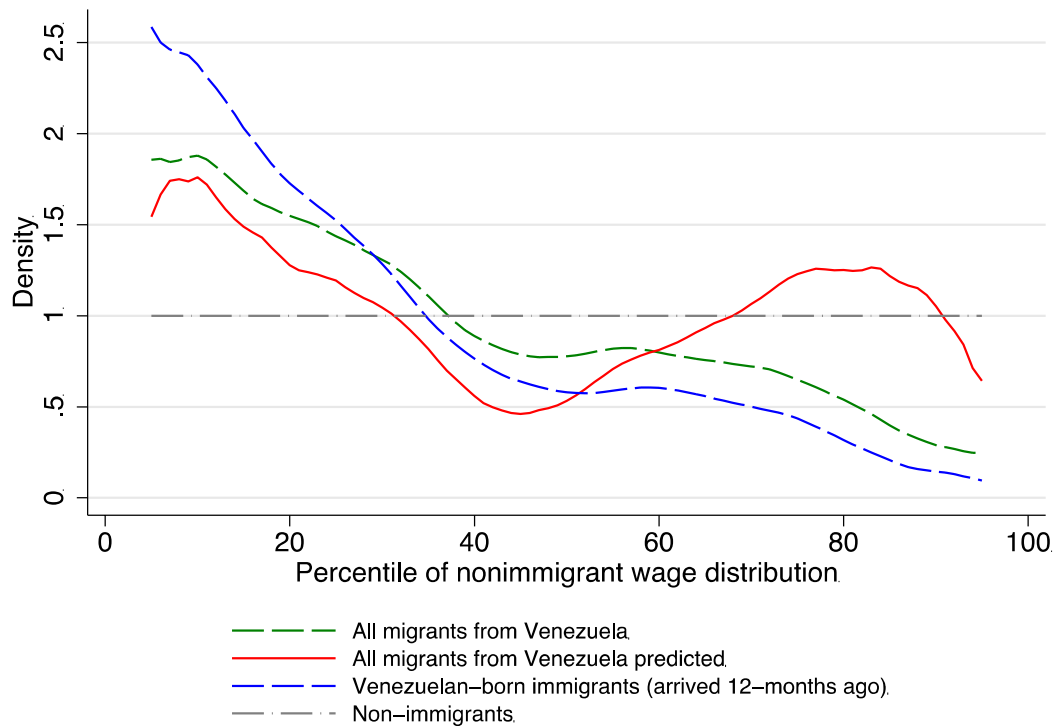
Notes: The Figure reports OLS and 2SLS coefficients and the respective 95% confidence intervals obtained by regressing natives' log wages, unemployment, employment, and labor force participation status on the fraction of immigrants (\tilde{m}_{jt}) separated by economic activity. IV1 instruments \tilde{m}_{jt} with the past settlement instrument as defined in Eq. (7). IV2 instruments \tilde{m}_{jt} with the distance instrument as defined in Eq. (9). All regressions include year and area fixed effects, individual controls (sex, age, age squared), and dummies for education achievement (less than high school, high school, some college, college graduates, and graduate degrees). Wages are computed for wage and salary workers and include the labor income of self-employed workers. The sample is restricted to natives aged 15 to 64 living in metropolitan areas (MSAs) from 2013-2018. To alleviate the potential impact of outliers, wages were computed by trimming the wage distribution by year at 0.5% and 99.5%. Wages are expressed in 2018 equivalent COP. Estimates are weighted by sampling weights. Standard errors are clustered at the metropolitan area level.

Figure 1.15 Effect of immigration on wages by type of job



Notes: The Figure reports OLS and 2SLS coefficients and the respective 95% confidence intervals obtained by regressing natives' log wages, unemployment, employment, and labor force participation status on the fraction of immigrants (\tilde{m}_{jt}) separated by type of job. IV1 instruments \tilde{m}_{jt} with the past settlement instrument as defined in Eq. (7). IV2 instruments \tilde{m}_{jt} with the distance instrument as defined in Eq. (9). All regressions include year and area fixed effects, individual controls (sex, age, age squared), and dummies for education achievement (less than high school, high school, some college, college graduates, and graduate degrees). Wages are computed for wage and salary workers and include the labor income of self-employed workers. The sample is restricted to natives aged 15 to 64 living in metropolitan areas (MSAs) from 2013-2018. To alleviate the potential impact of outliers, wages were computed by trimming the wage distribution by year at 0.5% and 99.5%. Wages are expressed in 2018 equivalent COP. Estimates are weighted by sampling weights. Standard errors are clustered at the metropolitan area level.

Figure 1.16 Actual vs. predicted position of Venezuelan immigrants in natives' wage distribution



Notes: The Figure depicts the observed kernel density (green dashed line) and the predicted kernel density estimates for all immigrants from Venezuela, including returnees and Venezuelan-born (red line), and for the Venezuelan-born immigrants that arrived in Colombia within the last 12 months. The predicted density is based on where Venezuelans would be located if they received the same return to education and experience as natives. The predicted line is obtained by estimating a flexible log regression separately for native males and females in each year between 2014 and 2018. The regression includes three age categories (15 to 28, 29 to 40, and 41 to 64), three education groups (less than high school diploma, high school diploma or technical degree, and bachelor's degree or more), the interaction between age and education groups, dummies for Bogotá, Medellín, and Cali, and dummies for each quarter of the survey. The resulting coefficients are then used to predict out-of-sample wages for different groups of immigrants from Venezuela. This procedure was based on Dustmann *et al.* (2013). The horizontal dashed gray line shows the native wage distribution as a reference. Kernel estimates are above (below) the horizontal gray line where migrants from Venezuela are more (less) concentrated than natives. Nonimmigrants are all natives, excluding returnees. The sample is restricted to the urban working-age population (15-64 years old) in the labor force who report labor income and are not enrolled in school. To alleviate the potential impact of outliers, wages were computed by trimming the wage distribution by year at 0.5% and 99.5%. Wages are expressed in 2018 Colombian pesos. *Source:* Own estimates using information from the GEIH 2014-2018 for all 23 MSAs.

Table 1.7 Robustness estimates of the effect of immigration on natives’ fiscal contributions – all budget

<i>Estimates of immigrant inflows (m_{jt})</i>	Coef.	SE	<i>F</i> -stat	Anderson–Rubin CI
<i>(1) Net fiscal contributions</i>				
(a) Using the shares from the 1993 census	0.008	0.033	23.707	[−0.046, 0.095]
(b) Normalizing ΔM_{jt} by the local pop. in 2013	0.007	0.035	22.672	[−0.050, 0.100]
(c) Using individual pooled data (\tilde{m}_{jt})	−0.012	0.014	25.567	[−0.040, 0.017]
(d) Using a distance-based instrument	−0.014	0.030	66.698	[−0.067, 0.049]
(e) Controlling for dynamic bias				
– Contemporaneous term: m_{jt}	0.191	0.101	30.433	[0.051, 0.612]
– Lagged term: m_{jt-1}	−0.278	0.121	74.634	[−0.699, −0.110]
(f) LIML	0.007	0.036	24.212	[−0.042, 0.097]
<i>(2) Revenues</i>				
(a) Using the shares from the 1993 census	0.076	0.037	23.707	[0.023, 0.187]
(b) Normalizing ΔM_{jt} by the local pop. in 2013	0.067	0.041	22.672	[0.001, 0.181]
(c) Using individual pooled data (\tilde{m}_{jt})	0.029	0.014	27.496	[0.003, 0.059]
(d) Using a distance-based instrument	0.036	0.038	66.698	[−0.032, 0.117]
(e) Controlling for dynamic bias				
– Contemporaneous term: m_{jt}	0.152	0.110	30.433	[−0.002, 0.535]
– Lagged term: m_{jt-1}	−0.125	0.139	74.634	[−0.510, 0.068]
(f) LIML	0.069	0.042	24.212	[0.014, 0.186]
<i>(3) Expenditures</i>				
(a) Using the shares from the 1993 census	0.051	0.024	23.707	[0.007, 0.106]
(b) Normalizing ΔM_{jt} by the local pop. in 2013	0.048	0.025	22.672	[−0.002, 0.105]
(c) Using individual pooled data (\tilde{m}_{jt})	0.034	0.007	25.935	[0.020, 0.047]
(d) Using a distance-based instrument	0.040	0.021	66.698	[0.003, 0.084]
(e) Controlling for dynamic bias				
– Contemporaneous term: m_{jt}	−0.070	0.074	30.433	[−0.275, 0.084]
– Lagged term: m_{jt-1}	0.179	0.073	74.634	[0.028, 0.381]
(f) LIML	0.049	0.026	24.212	[−0.002, 0.106]

Notes: The table reports various estimates of the effect of changes in the fraction of immigrants on natives’ net fiscal contributions, tax contributions, and expenditures. All regressions include year dummies and interactions of MSA-level controls with year dummies. Results are net of individual-level controls (sex, age, education) and computed by trimming the distribution of contributions each year at 1% and 99%. We report 5%-level identification-robust Anderson–Rubin confidence sets. Results are expressed as 2018 equivalent COP\$ million.

Table 1.8 Effect of immigration on natives’ outflows

Panel A. Variation across metropolitan areas				
	OLS	2SLS		
Immigrant share (\tilde{m}_{njt})	-0.007*** (0.001)	-0.010*** (0.001)		
Kleibergen-Paap F -stat		21.343		
Observations	138	138		
Panel B. Variation across neighborhoods (<i>tracts</i>)				
	OLS	OLS	2SLS	2SLS
Immigrant share (\tilde{m}_{njt})	0.012 (0.002)	0.017*** (0.002)	-0.008** (0.004)	-0.001 (0.019)
Area by Year FE		✓		✓
Kleibergen-Paap F -stat			43.982	27.413
Observations	8,186	8,186	8,186	8,186

Notes: The table reports the coefficients obtained by regressing the log of the number of natives on the fraction of immigrants (\tilde{m}_{kt}) between 2013–2018. Regressions in columns 1 and 3 include year and area (tract) fixed effects. Regressions in columns 2 and 4 control for tract and area by year fixed effects. The fraction of immigrants (\tilde{m}_{njt}) in Panel B is constructed by averaging immigrants in the tract using 2-year moving averages. In all regressions, observations are weighted by the total population in the area or tract in 2013. Columns 1 and 3 report in parentheses robust standard errors. Columns 2 and 4 report robust standard errors clustered at the area level. *** Denotes significance at 1%, ** significance at 5% and * significance at 10%.

B. Data Appendix

In this section we describe the allocation of revenues and expenditures to each individual observation in the GEIH. Before moving on, two clarifying points are in order. First, items in the budget can be assigned as individual-based or group-based revenues/expenditures. In other words, based on self-reported information in the GEIH we can attribute some tax contributions and benefits to each observation independently (individual-based). For those that we cannot attribute directly, we assign revenues and expenditures on a per capita basis using individual eligibility (group-based). Second, since we have fiscal information at all levels of government (National government, social security sector, and regional and local governments), all group-based revenues and expenditures are assigned conditional on the geographical location of the individual. Thus, when we are able, we separate the contribution of each municipality from the national contribution.

Tables 1.9 and 1.10 list all revenues and expenditures sources and their corresponding grouping.

B.1 Allocation of Revenues

B.1.1 Income tax, social insurance, and payroll taxes

Income tax contributions are estimated by applying year-specific tax rates of the ordinary system to gross annualized taxable income. We use

information from all sources of labor and capital income, non-labor compensation, the sale of fixed assets, and all payments for fees, commissions, services, emoluments, and among others reported in the GEIH.³⁶ We aggregate monthly information into annual income and apply the tax schedule of the ordinary system.³⁷ In our estimates, we allow for tax benefits that reduce the tax base: (i) costs and deductions, (ii) non-taxed income, and (iii) exempt income (25% for wage and salary workers and 50% for self-employed). In addition, we estimate tax withholdings for those individuals below the income threshold.³⁸

We estimate social security contributions as follow. Contributions to the contributory health care regime, Colpensiones, and the pension solidarity fund (FSP) are estimated using year-specific rates applied to the estimated Contribution Base Income (IBC) for all wage and salary workers and self-employed. Contributions to the minimum pension guarantee fund (FGPM) are estimated by taking 1.5% of the IBC for those enrolled in a private pension fund. Social security contributions to special regimes (*e.g.*, police and the national railway fund) are assigned per capita to all workers enrolled in a special regime.

³⁶ For 2017 and 2018, we estimate capital taxes including dividends as we cannot separately identify this in the GEIH.

³⁷ Before 2018, the tax each person had to pay was the highest value resulting from simultaneously declaring via the ordinary system and the National Alternative Minimum Tax (IMAN). For simplicity we use the ordinary system. This does not affect our results.

³⁸ According to Steiner and Cañas (2014), 76% of all income tax contributions in 2010 corresponded to tax withholdings of individuals who never filed a tax return.

Payroll taxes (SENA, ICBF, ESAP, and industrial schools) are estimated using year-specific rates applied to the estimated base salary.

B.1.2 Corporate and capital taxes

To assign corporate taxes, net of nonresident foreign ownership share, we first divide the contribution from national public ownership of companies and households. We allocate the national public ownership share on a per capita basis and assign the share corresponding to resident households equally to all those receiving individual dividend and interest income. Corporate taxes at the national level are distributed using the total population in the country, while those at the municipal level are only assigned to the residents of the municipality. As described in Mesa-Guerra & Ramírez-Tobón (2022), corporate and capital taxes include the share of other taxes (*e.g.*, wealth, indirect, and property taxes) that is paid by firms and not households.

B.1.3 Wealth tax

Wealth tax contributions of resident households are estimated by applying year-specific tax rates to net wealth. We use a multi-step procedure to estimate net wealth based on information from the GEIH. First, for property owners, we include self-reported values on the minimum price a household would sell their dwelling if they decided to do so. Next, we estimate price-to-rent ratios for each year using self-assessed housing

values and expected rent payments and apply these ratios to annualized self-reported individual rental income. We impute mean housing values by socio-economic strata to all missing values in each year and bottom-code to one million pesos. In our estimates, we allow for all tax deductions as recognized by law. The total tax attributed to each household is distributed among the household head and his or her spouse or partner using individual contributions to the combined household gross income.

B.1.4 VAT and other indirect taxes

We use estimates of decile-specific effective VAT tax rates from Mesa-Guerra & Ramírez-Tobón (2022) and apply these to the household gross income. We distribute total payments to indirect taxes to all members of the household using individual contributions to the household gross income.

B.1.5 Motor vehicle tax

We use a multi-step procedure to estimate motor vehicle tax contributions. First, we estimate the number of vehicles (cars and motorcycles) owned by the household. Second, we assign to each car the average value of modal car sold in Colombia (28 million pesos). We do the same for motorcycles using the 10 most sold brands (6.5 million pesos). Third, we apply the tax rate to each vehicle and use the resulting value to estimate each household's share in the total value aggregated at the municipal level.

Fourth, we take these shares to assign motor vehicle tax revenues to each household. Each household's contribution is split equally among all adult (18+) members.

B.1.6 Property tax

To estimate property tax contributions, we use self-assessed values of all real estate for property owners. Each household's share in the property tax of the corresponding municipality is distributed among the head of the household, his/her spouse or partner, and those members who report rental income and are classified as an independent person.³⁹ We impute the average for all missing observations in each socio-economic strata and bottom-code to one million pesos.

B.1.7 Industry and commerce tax

Assigned per capita to all self-employed owning an industrial, commercial, or service business in the municipality.

B.1.8 Financial transactions tax

³⁹ We consider as an independent person any individual age 18 years or older who is not enrolled full time in secondary education and if enrolled in higher education is working more than half time. If a person is married, they are considered independent, regardless of their age.

We use estimates of expenditures-to-income ratio by decile from Mesa-Guerra & Ramírez-Tobón (2022) and apply these to individual gross income. We then allocate financial transactions tax using an individual's expenditure share among those with monthly expenditures above 350 units of tax value.

B.1.9 Gross operating surplus, rents, and royalties

We assign gross operating surplus, rents, and royalties using per capita contributions to all adults (18+).

B.1.10 Urban phones tax

We assign urban phones tax using per capita contributions to all adults (18+) with access to a landline.

B.1.11 Educational services

We apportion revenues from educational services using per capita contributions to all individuals enrolled in higher education in public establishments.

B.1.12 National Teachers Pension Fund

We allocate these resources using per capita contributions to all workers in the educational sector (pre-school, primary, or secondary education) classified as wage and salary workers working for the government.

B.1.13 Other

All remaining tax payments, fees, fines and penalties, and other receipts, are assigned equally to all adults (18+).

Table 1.9 List of Government Revenues

Revenue source	Grouping
Personal income tax	Income tax, social insurance, and payroll taxes
Withholding tax	
Capital gains taxes	
Social security contributions	
Payroll taxes	
Corporate income tax	Corporate and capital taxes
Minimum presumptive tax	
Income tax for equality (CREE)	
Wealth tax	Wealth tax
Value-added tax	VAT and other indirect taxes
Consumption tax	
Wine and spirits duties	
Beer and cider duties	
Tobacco duties	
Carbon tax	
Fuel tax	
Fuel surcharge	
Tariffs and customs duties	
Vehicle excise duties	Motor vehicle tax
Property tax	Property tax
Industry and commerce tax	Industry and commerce tax
Financial transactions tax	Financial transactions tax
Gross operating surplus and rents	Gross operating surplus, rents and royalties
Interests and dividends	
Oil and mining royalties	
Urban phones tax	Urban phones tax
Educational services	Educational services
Social security contributions FOMAG	National Teachers Pension Fund
Other receipts FOMAG	
Fees and rights	Other
Fines and penalties	

Contributions
Sale of goods and services
Contractual income
Other taxes
Other receipts

Source. Mesa-Guerra & Ramírez-Tobón (2022).

B.2 Allocation of Expenditures

B.2.1 'Pure' and 'congestible' public goods

We allocate public goods on a per capita basis to the entire population.

B.2.2 Law courts and prisons

Law courts and prison expenditures are assigned per capita among the total adult (18+) population using the share of each origin-country group in the total prison population. Information on the origin-country of prison inmates comes from the National Penitentiary and Prison Institute (INPEC).

B.2.3 Water supply

Water supply expenditures are assigned equally to all households with access to water supply and then divided per capita among all household members.

B.2.4 Health services

To assign health services, we first identify all individuals enrolled in the General Health Care Social Security System (SGSSS) as reported in the GEIH. We then impute the per capita personal health care cost (UPC) by

age group and regime which is estimated by the Ministry of Health. We assign the complete value of the UPC to those affiliated to a defined contribution regime (including those affiliated to special regimes) or to the subsidized regime, and 45% of the subsidized UPC as the cost of health services for all individuals not enrolled in the SGSSS.⁴⁰

B.2.5 Education

Expenditures for education include expenditure for compulsory education, job training programs, and higher education. Use information on enrollment in the GEIH for each education level, we assign expenditures per capita to the relevant population. For compulsory education, we take all individuals aged 3 to 16 years attending compulsory education. For job training, we take all individuals enrolled in job training programs. For higher education, we take all individuals enrolled in higher education in public institutions.

B.2.6 Social protection

Social protection expenditures, both provided in the form of cash and in-kind benefits, includes benefits for sickness and disability, pension, family and children, unemployment, social housing, and vulnerable population. We assign directly all self-reported benefits received from government agencies in the GEIH, with the following caveats:

⁴⁰ This follows from Reina *et al.* (2018).

- (i) Sickness and disability: for those outside the labor force that left their job due to illness, we assign the per capita expenditure; for those that directly report receiving disability-related benefits (variable *p1661s4a1*), we assign the reported value or impute the average when the value is not reported. We include values that contain descriptions such as: “discapacitado”, “incapacidad”, “invalidéz”, “personas especiales”, or “cuidadores”.
- (ii) Pension: we assign values for pension income and Colombia Mayor (variables *p7500s2a1* and *p1661s3a1*). We include pension-related benefits not reported directly under pension income that contain descriptions such as: “Adulto Mayor”, “pension”, “BEPS”, “Colpensiones”, and “tercera edad” (variables *p1661s4a1* and *p1661s4a2*). We impute the minimum wage for pension income for all missing observations that reported to have received the benefit, and the average for Colombia Mayor.
- (iii) Family and children: we add all family and children benefits in the household and divide them evenly among all household members (variable *p1661s1a1*). We add income received from other government programs such as: “ICBF”, “Madres Comunitarias”, “Madres Sustitutas”, “Hogar Gestor”, or “Nutrición” (variables *p1661s4a1* and *p1661s4a2*). We impute the median income for all missing observations that reported to have received the benefit.
- (iv) Social housing: we add all housing-related benefits in the household and divide them evenly among all household members (variable *p7500s2a1* for 2013-2015; variable *p1661s4a2* for 2016-2018). We add income received from other government programs that contain descriptions such as: “vivienda”, “arriendo”, “casa”, “damnificado”, “desastre”, or “desalojo” (variables *p1661s4a1* and *p1661s4a2*). We impute the median income for all missing observations that reported to have received the benefit.
- (v) Vulnerable population: we add all self-reported benefits in the household and divide them evenly among all household members (variable *p1661s1a1*). We add income that contains descriptions such as: “alcaldia/municipio”, “gobernación/departamento”, “desplazado”, “victima”, “ayuda humanitaria”, “negritudes/afro”, “migracion”, “recicladores”, “reintegracion”, “reincertados”, “desmovilizados”, “familias en su tierra/guarda bosques”, “integración social”, or “jóvenes

en accion” (variables *p1661s4a1* and *p1661s4a2*). We impute the median income for all missing observations that reported to have received the benefit.

- (vi) Other: since some households report higher values for the total benefits received from the government than the value reported for each individual program, we create a variable that contains the remaining difference. We add this by household and divide it evenly among all members. We do this also in years when we cannot disaggregate benefits by program.

B.2.7 Debt service

We assign debt service on a per capita basis to the entire population after excluding interests for debt acquired as a result of immigrants’ arrival to the country. Using information on the year the debt was issued and the cohort of arrival of all immigrants, we exclude the share of immigrants in the total resident population by year of arrival. Debt service of regional and local governments is allocated per capita as we cannot disaggregate it by issuance date.

Table 1.10 List of Government Expenditures by Sub-Function (UN COFOG-DANE)

Expenditure item	Grouping
1.1 - 1.2 Executive and legislative organs, financial and fiscal affairs, external affairs, and foreign economic aid	
1.3 General services	
1.4 Basic research	
1.5 R&D general public administration	
1.6 General public administration n.e.c.	'Pure' public goods
1.8 Transfers of a general character between different levels of government	
2.1 Military defense	
2.2 Civil defense	
2.3 Foreign military aid	
2.4 R&D defense	
2.5 Defense administration	

3.1	Police services	
3.2	Fire-protection services	
3.5	R&D public order and safety	
3.6	Public order and safety administration	
4.1	General economic, commercial, and labor affairs	
4.2	Agriculture, forestry, fishing, and hunting	
4.3	Fuel and energy	
4.4	Mining, manufacturing, and construction	
4.5	Transport	
4.6	Communication	
4.7	Other industries	
4.8	R&D economic affairs	
4.9	Economic affairs n.e.c.	
5.1	Waste management	
5.2	Wastewater management	
5.3	Pollution abatement	
5.4	Protection of biodiversity and landscape	
5.5	R&D environmental protection	
5.6	Environmental protection administration	'Congestible' public goods
6.1	Housing development	
6.2	Community development	
6.4	Street lighting	
6.5	R&D housing and community amenities	
6.6	Housing and community amenities administration	
7.5	R&D health	
7.6	Health administration	
8.1	Recreational and sporting services	
8.2	Cultural services	
8.3	Broadcasting and publishing services	
8.4	Religious and other community services	
8.5	R&D recreation, culture, and religion	
8.6	Recreation, culture, and religion administration	
9.5	Education not definable by level	
9.6	Subsidiary services to education	
9.7	R&D education	
9.8	Education administration	
10.8	R&D social protection	
10.9	Social protection administration	
3.3	Law courts	
3.4	Prisons	Law courts and prisons
6.3	Water supply	Water supply
7.1 - 7.2 - 7.3	Medical, hospital and pharmaceutical services	
7.4	Public health services	Health services
9.1 - 9.2	Pre-primary, primary, and secondary education	Education: compulsory education
9.3	Post-secondary non-tertiary education	Education: job training
9.4	Tertiary education	Education: higher education
10.1	Sickness and disability	Social protection: sickness and disability

10.2 - 10.3	Old age	Social protection: pensions
10.4	Family and children	Social protection: family and children
10.5	Unemployment	Social protection: unemployment
10.6	Housing	Social protection: housing
10.7	Socially vulnerable and excluded population	Social protection: vulnerable population
11	Public debt service	Debt service

Source. Mesa-Guerra & Ramírez-Tobón (2022).

CHAPTER 2

IMMIGRATION AND NATIVES' INTERNAL MOBILITY IN THE DEVELOPING WORLD: EVIDENCE FROM THE VENEZUELAN IMMIGRATION TO COLOMBIA BETWEEN 2014-2019

2.1 Introduction

How natives' mobility decisions are affected by the arrival of immigration is central to understanding the labor market effects of immigration and the change in the demographic and ethnic composition of cities. Even though 35% of the migration is between developing countries, most of the literature on the impact of international migration supply shocks on internal mobility is concentrated among OECD countries (OECD, 2016). For the developed world, the evidence varied from moderately or nonexistent offsetting responses to international immigration in the US and France (Card, 2001; Card, 2007; Edo, 2020) to displacement effects in the cases of Germany, Italy, and Canada (Dustmann et al., 2017; Basile et al., 2021; Beine & Coulombe, 2018).⁴¹

⁴¹ In this document, I refer to displacement as natives moving out of areas experiencing higher immigration and crowd-out as natives reducing their inflows to areas experiencing higher immigration rates. The first relates to changes in natives' out-migration rates, whereas the former refers to changes in natives' in-migration rates. Although there is mixed evidence for the US, where some have found large crowd-out effects (Filer, 1992; Borjas et al., 1997; Hatton & Tani, 2005; Borjas, 2006) and others small and even positive evidence of native mobility responses to immigration (Card & Dinardo, 2000; Card, 2001; Card, 2007; Cortes, 2008), microsimulations studies of some empirical specifications used in the literature by Peri and Sparber (2011) tend to point in the direction that specifications as those defined by Card (2001) and Card 2007) seem better equipped to

This paper focuses on adding evidence of the effect of immigration on natives' internal migration decisions in the developing world. In particular, I use the mass inflow of Venezuelan immigration to Colombia and detailed labor-force data to study the effect of changes in immigration inflows across states on changes in natives' net migration rates and their components (in and out-migration rates) between 2014-2019. The arrival of Venezuelans into Colombia is one of the most explosive and recent examples of south-to-south migration or immigration between developing countries. As a result of the political instability and its socio-economic consequences, more than seven million Venezuelans have moved out of the country since 2015.⁴² Due to historical relations, cultural similarity, geographical proximity, and migration policies, Colombia became the leading destination for Venezuelans, accounting for almost 35% of the Venezuelan stock of immigrants as of 2019.⁴³⁴⁴ In some states in Colombia,

properly capture natives' location responses if present (at least in reduced-form specifications). More recently, however, evidence of large and even perfect crowd-out has been found for the US (Amior, 2021; Albert & Monras, 2020; Monras, 2020; Monras, 2021; Colas, 2018; Burstein et al., 2020; Abramitsky et al., 2019). Amior (2021) argues that omitted shocks play an important role and that estimates of crowd-out effects will be biased towards zero without the inclusion of proper controls. Including the initial employment rate can perform the function of partialling-out the entire history of immigration and demand shocks. Albert and Monras (2020) argue that internal migration is more responsive to local shocks in high-price locations. Burstein et al. (2020) find that immigrants crowd out natives in non-tradable occupations.

⁴² <https://www.imf.org/en/News/Articles/2022/12/06/cf-venezuelas-migrants-bring-economic-opportunity-to-latin-america>

⁴³ Venezuela's humanitarian crisis triggered by the food shortages, limitations in access to health care, and Bolivar's devaluation created unprecedented migration flows across Latin America's countries, shifting the migration profile of Venezuela from a country of destination to a country of origin. As of September 2018, Venezuela experienced an

immigrants from Venezuela went from representing less than 1% of the working-age population in 2014 to more than 6% in 2018 (see Figure 2.5 in the Appendix).⁴⁵

Methodologically, I follow Peri & Sparber (2011) and define specifications in the spirit of Card (2001, 2007), connecting year-to-year changes in immigrant population to changes in natives' net migration rates and its components (in and out-migration rates) across 23 Colombian states over the period 2014-2019. The preferred empirical specification uses variation in migration inflows across states and over time, and an instrumental variable strategy whereby changes in immigration population are instrumented using two Bartik shift-share instruments: the enclave instrument based on previous Venezuelan settlements in Colombia (similar in spirit to Altonji & Card, 1991; Card, 2001; Card, 2007), and a distance-based instrument (similar to Del Carpio & Wagner, 2015).

I found that although OLS and IV estimates suggest a negative correlation between aggregated changes in immigration from Venezuela and natives' internal migration decisions, once confidence sets robust to the presence of weak instruments are constructed, the effect is not

inflation rate of 500,000%, and the associated decrease in real wages has caused 90% of the population to live in poverty (World Bank, 2018).

⁴⁴ As of June 2019, 1.4 million Venezuelans lived in Colombia:

<http://www.migracioncolombia.gov.co/index.php/es/prensa/infografias/infografias-2019/12565-infografia-venezolanos-en-colombia>

distinguishable from zero. Similarly, changes in immigration inflows from Venezuela are positively associated with natives' out-migration and negatively associated with natives' in-migration rates in models controlling for state and year fixed effects. However, effects for both natives' in and out-migration rates are indistinguishable from zero once robust confidence sets are constructed. Results imply that although the direction of the effects seems consistent with the idea that natives are inclined to leave states receiving high inflows of immigrants from Venezuela (consistent with displacement) or to avoid migration into those states (consistent with crowding out), I am not able to find a statistically significant relationship in our sample of states for the period 2014-2019.

Whether developing countries receiving a mass influx of immigrants respond similarly or differently to developed countries and under what conditions this is the case remains an open question. Results are consistent with previous findings using similar empirical specifications for the developed world, where natives' location responses are moderate to nonexistent (Card, 2001; Card, 2007; Cortes, 2008; Edo, 2020). Moreover, the document complements the growing body of literature on native responses to immigration by adding evidence from one of the most prominent examples of south-to-south migration over the last 20 years: the arrival of immigrants from Venezuela to Colombia. Few studies have systematically studied the effect of immigration shocks on internal

migration in the developing world. Del Carpio et al. (2015) studied Malaysians' adjustments to immigration using detailed Labor Force Survey data between 1990-2010. They found that immigration induced positive inflows of natives. For every ten immigrants into a given state, there was an additional inflow of almost eight working-age natives, with most of the effect driven by natives finding new job opportunities and less by natives being out of the labor force.

My results stand in contrast to the Malaysian experience despite the immigrants' profiles being relatively similar. Although Venezuelans arriving in Colombia are, on average, more educated than Colombians, a substantial percentage of them remained in the country irregularly and therefore were potentially competing with natives in the lower part of the skill distribution. The immigration to Malaysia was comparatively similar, with most immigrants being less educated than Malaysians and a high percentage of immigrants having to stay in the country irregularly due to stringent immigration policies where work permits were given only for up to a year.

There are several reasons why results for two developing countries experiencing immigration shocks differ. First, Del Carpio et al. (2015) also explore margins of adjustment outside the working-age population. They explore how native labor supply responds to immigration by continuing

education, retiring, or being more or less likely to be employed or unemployed. I focus on working-age natives' location decisions in response to changes in the intensity of immigration. Second, natives could respond to immigration shocks over a longer horizon. Therefore, the two decades of data used for the Malaysian case could have facilitated capture of natives' responses, in contrast to the Colombian case, where I only used variation between 2014-2019.

Third, results could have also differed due to the specific differences in the labor markets receiving each shock, specifically regarding housing prices. Albert & Monras (2020) show that internal migration is more responsive in high than low housing price locations. There may have been no substantial differences in housing prices between cities attracting immigrants in Malaysia, which could have been the case in Colombia. However, in the Colombian case, it is challenging to connect housing prices to changes in natives' location response because the indexes measuring changes in housing prices in Colombia are built for 13 main metropolitan areas and not for states which are our geographic unit of analysis.

The document also contributes to the validity of findings in the growing body of literature studying the labor markets effects of Venezuelan immigration across Colombian local labor markets (Bonilla et al., 2023; Lebow, 2022; Delgado-Prieto, 2022; Caruso et al., 2019). If

natives respond rapidly to the arrival of Venezuelans to a given area, then estimates that do not account for those movements could underestimate the effect of immigration on natives' wages and other labor market outcomes. The evidence about the impact of Venezuelan immigration on natives' location decisions is mixed. Bonilla et al. (2023) show that immigration reduces the probability of in-migration, whereas Caruso et al. (2019) and Lebow (2022) find no statistically significant effects on either in-migration or out-migration.⁴⁶

However, I argue that empirical specifications and variable constructions used in previous literature for Colombia are limited and do not allow to appropriately capture natives' location response to one additional immigrant arriving at a given location. First, specifications are not well equipped to capture displacement or crowding out effects. Bonilla et al. (2023) and Caruso et al. (2019) define specifications that mechanically induce a correlation between migration changes and native location responses because the counts of natives are both used to build the

⁴⁶ Bonilla et al., 2023 run models where the outcomes variable, immigration and outmigration, are measured as the log of the number of individuals who arrive or left a Colombian state during the last 12 months (excluding all migrants from Venezuela). They found that immigration reduces natives' in-migration, implying that immigrants crowd out natives from areas with high immigration. Caruso et al. (2019) ran LPM for in-migration and out-migration. They found a negative correlation between immigration shares (defined as the fraction of Venezuelan immigrants in the working-age population) and natives' in-migration and positively correlated to immigrants. However, the effects were not statistically significant. Lebow (2022) ran a first-difference specification between 2014-2019 relating the effect of Venezuelan arrivals on the normalized count of natives who entered (in migration) or left (out-migration) a metro area over the previous five years as a share of the 5-year lagged population. He found similar results to Caruso et al. (2019).

outcome variable and the policy variable, the fraction of immigrants in a given labor market (this problem is discussed in Peri & Sparber, 2011).⁴⁷ All three documents fail to appropriately normalize the variables to facilitate interpretation and to be able to capture the native response to one additional immigrant. Lebow's (2022) specification has some limitations as it only uses changes in internal migration over five years. However, the empirical specifications defined to study the labor market effects exploited yearly variation across states.⁴⁸

I contribute to this literature by connecting natives' location responses to immigration from Venezuela through empirical specifications well designed to identify native location responses (as discussed by Peri & Sparber, 2011) and by also exploiting year-to-year variation across geographic units as has been standard practice in most of the documents interested in the labor market effects of Venezuelan immigration in Colombia. Contrary to Bonilla et al. (2023), but in line with Caruso et al. (2019) and Lebow (2022), I find no offsetting native location responses to

⁴⁷ Both Bonilla et al (2023) and Caruso et al (2019) define the migration shares as (working-age immigrants / working-age natives), whereas the outcome is defined as the number of natives that move in or out a given state over the last 12 months. It's highly likely that many individuals in the working age population (on the left side of the equation) are also internal migrant counts (on the right side of the equation), inducing a mechanical relationship between the variables.

⁴⁸ If native location decisions are happening more rapidly and the adjustment is quicker (over a few months or a year) as some evidence suggests (Monras, 2020), then specifications not capturing those years to year variations in natives' responses to immigration could paint the picture of non-existent native responses because the time horizon of analysis and variation used for wage or other labor market outcomes is different from the time horizon and variation used for understanding natives' response to immigration.

Venezuelan immigration over the period 2014-2019. Results are important to confirm that previous findings on the effects of Venezuelan immigration on natives' wages (especially for those less educated) were not biased due to natives' moving out of or avoiding places with a higher concentration of immigrants from Venezuela.

Finally, the document contributes to the literature on population responses to supply and demand shocks in Colombia. Previous literature has focused on the response to demand shocks (Baron, 2011) and the role of forced displacement supply shocks on various economic outcomes. Previous findings show several margins of adjustment. Calderon-Mejia & Ibanez (2016) shows that forced displacement supply shocks reduce wages for urban low-skilled workers. Morales (2017) found persistent negative effects on wages for low-skilled women. Boozoli et al. (2012) found that forced displacement increases self-employment, mainly in the service sector. I add to this literature by showing evidence that natives' location responses to supply shocks are less sensitive than to demand shocks, as reported by Baron (2011), who shows that natives tend to migrate to cities experiencing better economic prospects.

The rest of the document is organized as follows. Section 2 presents a simple accounting framework connecting changes in immigration to changes in the native population to guide our empirical specifications.

Section 3 describes the data. Section 4 presents the empirical strategy and discusses the identification assumptions. Section 5 presents and discusses the reduced form evidence on the effect of Venezuelan immigration on natives' net migration rates and its components and robustness checks. Section 6 concludes.

2.2 Theoretical Framework

I use a simple model of population growth to guide our empirical specifications. The idea of the model is to connect the arrival of immigration from Venezuela to changes in local populations. At the national level, the arrival of one additional immigrant adds one additional member to the total population. However, at the local level, the arrival of immigrants could be offset by natives deciding to change the location for economic and demographic considerations or in response to the arrival of immigrants.

Let L_t be the total working-age population in a given state at a time t . L_t is the sum of native-born people in the state at that time (N_t) and the number of immigrants that arrived from Venezuela (F_t). Therefore, $L_t = N_t + F_t$. I can decompose the overall growth rate for a given state into the sum of changes in the native and the Venezuelan immigrant populations, divided by the initial total working age population in the preceding period $t - 1$ as:

$$\left(\frac{L_t - L_{t-1}}{L_{t-1}}\right) = \left(\frac{N_t - N_{t-1}}{L_{t-1}}\right) + \left(\frac{F_t - F_{t-1}}{L_{t-1}}\right) \quad (1)$$

Using the accounting equation (1), I can then define an empirical specification connecting changes in immigration from Venezuela given by $\left(\frac{F_t - F_{t-1}}{L_{t-1}}\right)$ to natives' population growth across states $\left(\frac{N_t - N_{t-1}}{L_{t-1}}\right)$. A similar setup has been used in the literature. Card (2007) used to test the effects of the arrival of foreign-born immigrants on changes in population growth across a single cross-section of US cities between 1980 and 2000. Since I define two mutually exclusive groups and use questions about internal migration decisions over the last 12 months, I can capture changes in the native population across states related to changes in net internal migration and its components (in and out-migration). In section 4, I present a causal model connecting changes in immigration from Venezuela to natives' population growth based on equation (1).

2.3 Data and Descriptive Statistics

The primary dataset used in this paper comes from Colombia's Labor Force Survey (Gran Encuesta Integrada de Hogares - GEIH) for 2013-2019. The GEIH is a nationally representative household survey that collects information on labor market conditions. Since 2009, the sample size has been 20,700 households every month. For population purposes, including the immigrant population, the survey is representative at the state level

for 23 Colombian states. This survey has three crucial advantages. First, it contains information on the geographic distribution of immigrants from Venezuela across states, allowing us to use empirical specifications that rely on variation across geographic units.⁴⁹ Second, immigrants are surveyed regardless of their migratory status (regular or irregular).⁵⁰ Third, the survey also captures interregional migration patterns since it asks questions about a 12-month and 5-year previous state or municipality of residence. This allows us to compute net migration and in-and-out-migration rates for the native population at the state level and connect changes in the arrival of immigrants to natives leaving or arriving in states.

I begin by aggregating the monthly dataset to the year and state levels for 2013-2019. For the empirical analysis, I restrict the sample to the working-age population (18 and 64 years old), who are not enrolled in school, and have a least one year of potential experience.⁵¹ The sample excludes migrants from other countries, which have been very stable

⁴⁹ The states used in the empirical analysis are Antioquia, Atlántico, Bogotá, Bolívar, Boyacá, Caldas, Caquetá, Cauca, Cesar, Córdoba, Chocó, Huila, La Guajira, Magdalena, Meta, Nariño, Norte de Santander, Quindío, Risaralda, Santander, Sucre, Tolima, and Valle del Cauca.

⁵⁰ I follow the terminology used by the Colombian government to classify immigrants. Undocumented or irregular immigrants do not satisfy the requirements established by the host country to enter or remain in the country.

⁵¹ Potential experience is defined as age minus age of workforce entry. Age of workforce entry is imputed based on education. I assume that those with less than high-school education enter the workforce at 17, those with high-school at 19, those with some college at 21, and those with college or more at 23.

throughout the study period (representing around 0.2% of the working-age population)⁵², and natives arriving in the country from abroad over the last 12 months. Internal migration decisions for the latter group are less sensitive to changes in the immigration inflows into a region and more related to family ties in those regions, so the groups excluded are remarkably stable across time.

Based on the model of population growth presented in section 2 at the state level, I define two mutually exclusive groups: 1) Natives, which correspond to all-native born, excluding returnees, and 2) Venezuelan *immigrants* as all individuals born in Venezuela plus all individuals born in Colombia who lived in Venezuela for 1 or 5 years (also known as returnees) before being surveyed. I use the 12-month previous residence to capture yearly native population in-flows, out-flows, and net migration flows for each state. This variable helps us to construct the net migration and in and out-migration rates for the native population at the state level for each year between 2013 and 2019. To facilitate interpretation and avoid spurious correlation between changes in immigration population and native internal migration rates, I standardize native and Venezuelan

⁵² Figure A1 in the appendix shows the evolution of the different migrant groups. Immigration from countries other than Venezuela has remained stable over the period 2013-2019, and therefore, it's not expected to contribute to changes in native population across states. Figure A2 in the appendix shows that after the opening of the border between the two countries in July 2016, the arrival of immigrants from Venezuela grew exponentially. As of 2019, In some states, Norte de Santander and La Guajira, immigrants from Venezuela represent more than 15% of the native working-age population.

immigrants' populations by the working age population at a base year (2013) and multiply it by 100.

Table 2.1 presents descriptive statistics for natives and immigrants from Venezuela for the first and last year of the period of study (Panel A) and descriptive statistics for the variables used in the aggregated regressions as presented later in equations (3) and (4).

Table 2.1 Descriptive statistics

<i>Panel A: Demographics</i>				
	Natives		Immigrants from Venezuela	
	2014	2019	2014	2019
Share population	0.995	0.949	0.005	0.051
Mean Age	39.7	40.1	34.6	33.3
Mean Schooling	8.7	9.3	8.8	9.9
Fraction College degree or more	0.10	0.11	0.08	0.10
Fraction manufacturing	0.12	0.12	0.15	0.11
Fraction Commerce	0.20	0.20	0.18	0.23
Fraction informal	0.62	0.59	0.65	0.72
<i>Panel B: Descriptive statistics variables used in aggregated estimations 2014-2019</i>				
	Mean	SD	Max	Min
Total Population Growth	0.89	1.41	7.28	-2.31
Venezuelan Immigration Population Growth	0.90	1.13	6.79	-0.21
Native Population Growth	-0.02	0.99	5.42	-2.45
Native IN - Migration Rate	4.02	1.87	14.19	1.06
Native OUT - Migration Rate	4.04	1.26	9.86	2.31
<i>Controls</i>				
Natives Mean Age	39.8	12.5	64	18
Natives Mean Age Squared	1,744	1,031	4,096	324
Natives Mean Potential Experience	21.2	12.9	47	1
Natives Mean Potential experience squared	615.2	599.3	2,209	1
Natives fraction college degree or more	0.13	0.33	1	0
Natives fraction in manufacturing	0.12	0.32	1	0
Natives fraction informality	0.61	0.49	1	0

Notes: Panel A reports for 2014 and 2019 descriptive statistics for natives and immigrants from Venezuela (including Venezuelan-born and returnees) across 23 Colombian states. Observations in panel A are weighted using the sampling weights from the 2005 Census. Panel B reports descriptive statistics for the variables used in regressions using aggregated data from 2014-2019. The natives' population growth is defined as $(WorkingAgePop_t - WorkingAgePop_{t-1}/WorkingAgePop_{2013})$. The net native migration rate is defined per 100 of the working-age population in 2013 $(InMigration - OutMigration/WorkingAgePop_{2013}) * 100$. Aggregated rates presented in rows (1-5) in panel B are weighted using the working-age population for each state in 2013. Informality is

defined according to the Colombian Bureau of Statistics (DANE). The sample was restricted to individuals aged 18 to 64, not enrolled in school, and with at least one year of potential experience. Source: Own calculations using data from the GEIH, 2014-2019.

Over the span of five years, there was a marked increase in the contribution of immigrants from Venezuela to the working-age population in the country. Immigrants from Venezuela went from representing 0.5% of the working-age population in 2014 to 5.1% in 2019. Figure 2.5 in the appendix shows that after the opening of the border between the two countries in July 2016, the arrival of immigrants from Venezuela grew exponentially. As of 2019, in some states, Norte de Santander and La Guajira, immigrants from Venezuela represent more than 15% of the native working-age population.

Immigrants from Venezuela are also younger and more educated on average than natives. As of 2019, the mean age for natives is 40, and for immigrants from Venezuela is 33 years old. The mean years of schooling for natives is 9.3, whereas for immigrants from Venezuela is 9.9. The demographic profile of immigrants from Venezuela has also changed over the period 2014-2019. Immigrants arriving in the country after 2017 are, on average, younger than those who came before 2015.

Are changes in native population related to changes in immigration from Venezuela across states? Figure 2.1 explores this relationship for two years, 2017 and 2019. Changes in the native population across Colombian states are negatively correlated to changes in the arrival of new

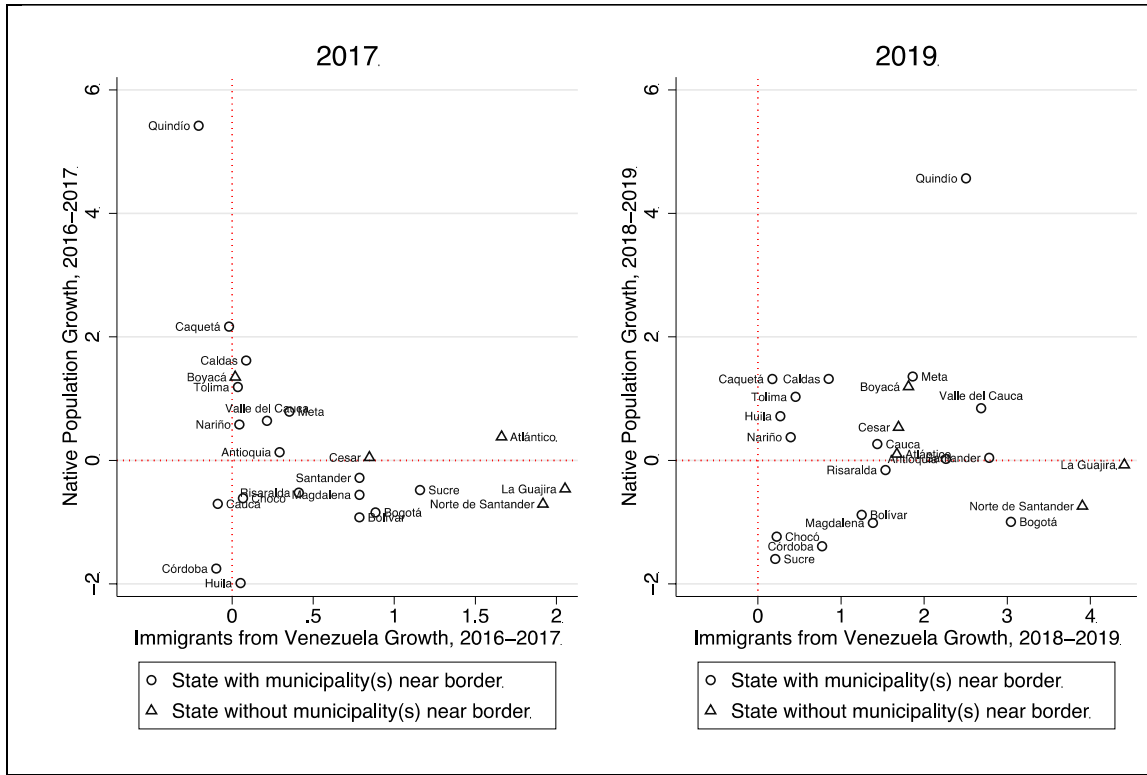
immigrants from Venezuela. The correlation coefficient for this relationship is -0.32 and 0.06 for 2017 and 2019, respectively.

The correlation coefficient between the two variables over the 2014-2019 period is -0.03 . In the initial years of our sample (2014 and 2015), the relationship between the variables was strong, but with changes in direction from 0.26 in 2014 to -0.35 . However, once the migration flows started to grow exponentially due to the reopening of the border between the two countries at the end of 2016, changes in immigration from Venezuela became an important contributor to changes in the native population across the Colombian state with a correlation of -0.20 between 2016-2018.⁵³

Changes in net migration rates explain the changes in native populations across states. The natives' net migration for a given state in a given year is defined as the difference between migration into the state (in-migration) minus migration out of the state (out-migration) divided by the state's working age population in 2013. This number is multiplied by 100 so that the net migration rate represents the net native internal migrants arriving in the state over the last 12 months per 100 of the working population in a base year (2013).

⁵³ In August 2015, the Venezuelan government unilaterally decided to close the border between the two countries and expel many Colombians living in the border regions. The border was reopened on August 12, 2016, dramatically increasing immigration from Venezuela.

Figure 2.1 Changes in native population vs. changes in Venezuela immigration, 2017 and 2019



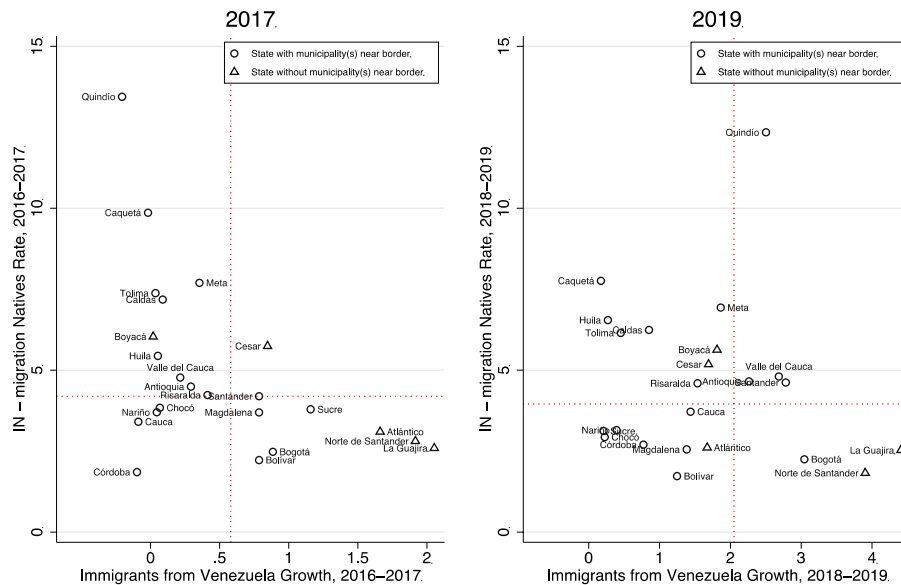
Notes: The Figure depicts the relationship between the normalized native population growth $((NativePop_t - NativePop_{t-1})/Pop_{2013}) * 100$ and the normalized Venezuelan immigration population growth rate $((VenzPop_t - VenzPop_{t-1})/Pop_{2013}) * 100$ across 23 states for the years 2017 and 2019. Since changes in internal migration explain the population growth, the figure is similar to depicting natives' net migration rates against changes in population growth from Venezuela. The net native migration rate is defined per 100 of the working-age population in 2014 $(InMigration - OutMigration/WorkingAgePop_{2013}) * 100$. The in-migration is the number of new natives arriving in the state over the last 12 months. The out-migration is the number of natives leaving the state over the last 12 months. Triangle points represent states with one or more municipalities at the border between Colombia and Venezuela. Source: Own calculations using data from the GEIH 2013-2019.

What drives the changes in net-migration rates? Figure 2.2 shows the relationship between the two components of the net migration rate, out-migration and in-migration rates, and the changes in immigration from Venezuela. There is a negative correlation between both in and out-migration rates and the changes in immigrants from Venezuela across states. This means, at least for the years depicted in the figure, that

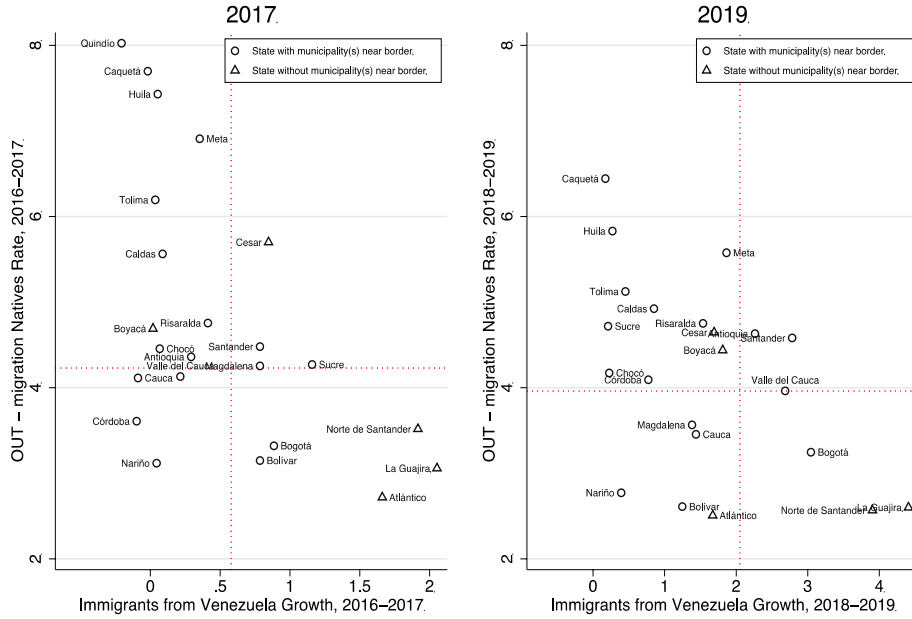
natives are not leaving states where immigrants are arriving. Instead, they are avoiding states where the population of Venezuelans is growing disproportionately. The descriptive relationships presented in figures 2.1 and 2.2, along with the accounting identity in equation (1), motivate my empirical approach in the next section.

Figure 2.2 Changes in natives' in and out migrations rates vs. Changes in Immigration from Venezuela, 2017 and 2019

Panel A. Natives In-migration



Panel B. Natives Out-migration



Notes: The Figure depicts the relationship between the natives' in-migration rates (Panel A) and out-migration rates (Panel B) and the changes in immigration from Venezuelans across 23 states for 2017 and 2019. Triangle points represent states with one or more municipalities at the border between Colombia and Venezuela. Source: Own calculations using data from the GEIH.

2.4 Empirical Strategy

Equation (1) can be translated into a causal equation connecting changes in immigration to changes in native populations across states. I follow Card (2007) and Peri & Sparber (2011) and define the empirical specification as follows:

$$\left(\frac{L_{c,t} - L_{c,t-1}}{L_{c,2013}}\right) = \alpha + \beta \left(\frac{F_{c,t} - F_{c,t-1}}{L_{c,2013}}\right) + \gamma X_{c,t} + \Omega_t + \Omega_c + \varepsilon_{c,t} \quad (2)$$

where $L_{c,t}$ is the population of state c in year t . $L_{c,t-1}$ is the population of state c in the previous year. $X_{c,t}$ is a vector of states' characteristics that can affect natives' location decisions. Those characteristics include average

demographics for natives (mean age, mean experience, and their squares), variables capturing industrial structure (the fraction of natives in manufacturing), and variables capturing economic conditions (fraction of informal employment).⁵⁴ The controls included are similar to those included by Card (2001). Ω_t and Ω_c are year and state-fixed effects, respectively. Ω_c could capture permanent (time-unvarying) differences in productivity across cities or unobserved state-level factors (such as local demand shocks) that might be correlated with immigrant inflows and native migration flows. Ω_t are meant to pick up overall trends similarly affecting all the states.

The coefficient β captures the causal effect of immigration inflows on overall population growth across Colombian states. For instance, $\beta = 1$ implies that one additional immigrant adds 1 to the overall population. We would observe native mobility responses if $\beta < 1$. For example, a value of less than one means that for each additional ten immigrants, a total of $10(1 - \beta)$ native-born residents would leave the state. Peri & Sparber (2011) suggest replacing $\left(\frac{L_{c,t} - L_{c,t-1}}{L_{c,2013}}\right)$ for $\left(\frac{N_{c,t} - N_{c,t-1}}{L_{c,2013}}\right)$ in equation (2) such that the empirical specification can be defined as:

⁵⁴ As a robustness exercise, I also ran a similar specification with state controls fixed in a baseline year before the shock (2013) to avoid potential pitfalls from controlling for intermediate outcomes that can also be affected by the shock. We get comparatively similar results.

$$\left(\frac{N_{c,t} - N_{c,t-1}}{L_{c,2013}}\right) = \alpha + \beta \left(\frac{F_{c,t} - F_{c,t-1}}{L_{c,2013}}\right) + \gamma X_{c,t} + \Omega_t + \Omega_c + \varepsilon_{c,t} \quad (3)$$

Where $N_{c,t}$ is the native population of state c in year t . $N_{c,t-1}$ is the native population of state c in the previous year. Based on microsimulation exercises from Peri & Sparber (2011), specifications such as equation 3 seem to capture the native displacement effect appropriately. In this specification, displacement is observed if $\beta < 0$.⁵⁵ The specification has several advantages. First, it does provide a more direct test for the displacement effect since we are connecting changes in immigration inflows to changes in the native population more directly. Second, since the specification is normalized by the local population in a base year ($L_{c,2013}$), it would not force a spurious correlation between the inflow of immigrants from Venezuela and the native population response to the arrival of immigrants (as it does in specifications such as those from Bonilla et al. (2023), Lebow (2022), and Caruso (2019)). The specification also includes the additional controls as in equation 2. The inclusion of Ω_c means that the effect of changes in immigration is identified from changes in normalized net-migration rates over time within states. The inclusion of Ω_t means that

⁵⁵ Since Peri & Sparber (2011) used decennial data between 1970 and 2010 and defined groups by state, year, and skill to test what specifications in the literature could adequately predict displacement, their specification included year, skill, region, and interactions year by region, year by skill, and skill by region fixed effects. My goal is to estimate the causal impact of changes in the inflow of migrants from Venezuela on natives' location decisions. Therefore, I also control for potential confounders by including local characteristics.

only deviations from year-specific national averages are used for identification.

To get a better understanding of what drives changes in population growth across states in response to changes in immigration inflows from Venezuela, I also define a specification connecting the different components of natives' net internal migration rates (in and out-migration rates) to year-to-year changes in immigrant inflows from Venezuela as follows:

$$\left(\frac{IM_{c,t}}{L_{c,2013}}\right) = \alpha + \beta \left(\frac{F_{c,t} - F_{c,t-1}}{L_{c,2013}}\right) + \gamma X_{c,t} + \Omega_t + \Omega_c + \varepsilon_{c,t} \quad (4)$$

Where $IM_{c,t}$ are counts of each one of the components of the net migration rates into state c over the last 12 months. Meaning counts of those joining or leaving state c over the last 12 months. Since population changes observed in a given state between two consecutive years are explained by differences in net internal migration, the estimated coefficients β for each component of the net migration rate estimated by equation (4) would add up to the coefficient β obtained by estimating equation (3).

However, some threats to the identification of coefficient β remain. Immigrants could be selecting themselves into areas experiencing positive demand shocks and, therefore, also attracting the native population into

those areas. Although state-specific fixed effects could capture any unobserved time unvarying characteristic of a particular state that leads to greater population inflows or lower out-flows (location or weather), unobserved state-specific factors might still present, such as productivity shocks ($\varepsilon_{c,t}$) that could be correlated with changes in immigration inflows into a state (or contributing to slower natives' out-flows), causing the coefficient β to be upward biased.

In addition, because labor force surveys are not initially intended to be a representative sample of all immigrant populations but of the overall population, this could lead to measurement error in the counts of immigrants and, therefore, to potential attenuation bias when trying to capture the effect of immigration on changes in natives' population. Although this has been discussed when trying to estimate the effect of immigration on natives' wages and other labor market outcomes (see, for instance, Aydemir & Borjas, 2011), the implications remain similar for understanding the effect of changes in immigration inflows on natives' location decisions.

To deal with potential endogeneity issues, I follow the literature and instrument $\frac{F_{c,t}-F_{c,t-1}}{L_{c,2014}}$ in equation (4) with two instruments.⁵⁶ First, I use

⁵⁶ The use of instrumental variables could also account for the attenuation bias generated by the fact that for some states in early years, the number of Venezuelan immigrants is low and, in some cases, even zero. This will be the case as long as the measurement error

historical settlement patterns of Venezuelan immigrants in Colombia. The instrument takes advantage of the well-known fact that new immigrants tend to locate near earlier immigrants from the same origin country, which, coupled with the annual variation of Venezuelans arriving in Colombia due to political and economic conditions in Venezuela, creates a plausible exogenous prediction of the immigration shock across spatial units (Altonji & Card, 1991; Card, 2001; Card, 2007). I use the 2005 Population Census to construct the geographical distribution of Venezuelans in Colombia well before the immigration shock happened. I construct this instrument of predicted inflows to each area as a fraction of the corresponding local population in 2013 as follows:⁵⁷

$$IV_{ct}^{Past} = \left(\frac{M_{c,2005}}{M_{2005}} \right) \times \left(\frac{1}{L_{c,2013}} \right) \times \Delta M_t. \quad (5)$$

Here, M_{2005} is the number of Venezuelan-born immigrants in Colombia in 2005, $M_{c,2005}$ is the number of Venezuelan-born immigrants in state c in 2005, $L_{c,2013}$ is the working-age population of state c in 2013, and ΔM_t is the number of working-age individuals arriving in Colombia each year from Venezuela. Notice that the share $\left(\frac{M_{c,2005}}{M_{2005}} \right)$ includes all Venezuelan-

in the selected instrumental variables is uncorrelated with measurement error in the spatial distribution of Venezuelans for the period 2013-2019. Since the chosen instrumental variables come from the 2005 Population Census, its potential measurement error is likely independent of the potential measurement error in the labor force survey that comes from a different dataset (GEIH).

⁵⁷ As a robustness check, I also construct the same instrument using the 1993 population census. I present results robust to the use of this instrument in section 5.2 of the results.

born immigrants in Colombia, while the shift (ΔM_t), restricts inflows to all working-age Venezuelans arriving to the country.

Identification comes from exogenous exposure (given by the shares) to a common shock (given by the shift), as discussed by Goldsmith-Pinkham *et al.* (2020). The validity of the instrument depends on the exogenous variation in the national inflow rates from Venezuelan to local conditions in a specific city. This condition is likely to hold because increasing inflows of Venezuelan immigrants during our analysis period are driven primarily by political and economic factors in Venezuela that are uncorrelated with specific economic or political conditions in Colombia.⁵⁸

The second is a distance-based instrument in the spirit of Del Carpio & Wagner (2015). The instrument has also been used to study the impact of immigration from Venezuelan on natives' wages (Caruso *et al.*, 2019; Delgado-Prieto, 2022). The instrument predicts the number of immigrants that arrived in each department c in year t as follows:

$$IV_{ct}^{Dist} = \sum_v \frac{1}{T_{vc}} \theta_v \cdot \Delta M_t, \quad (6)$$

⁵⁸ This assumption could be plausible considering that, as shown in Appendix A1, immigration from other countries did not significantly change between 2014 and 2019. If economic and political conditions in Colombia were driving the arrival of Venezuelans, then something similar should be observed for migrants from other countries.

where T_{vc} is the road distance between the centroid of the capital city in each Venezuelan province v and the centroid of the capital city in each Colombian state c .⁵⁹ θ_v is the distribution of Venezuelan-born immigrants by region of origin in Venezuela from the voluntary registry of undocumented migrants (Administrative Register of Migrants from Venezuela—RAMV) conducted in 2018.⁶⁰ ΔM_t corresponds to the new inflows of Venezuelan immigrants to Colombia in year t .⁶¹

Considering that Colombia and Venezuela share a border of more than 1,250 miles, the instrument exploits the fact that distance implies time and economic costs for new immigrants. Therefore, it is a significant contributing factor for Venezuelans considering migrating to Colombia.⁶² In regression using the distance-based instrument, identification comes from comparing cities in areas close to the border with those further away.⁶³ A major threat to identification with this instrument is that cities

⁵⁹ The distance is calculated using the package `georoute` in Stata 17 (Weber & Péclat, 2017).

⁶⁰ The information can be accessed here: <https://data.unhcr.org/en/documents/details/64101>

⁶¹ The instrument has a limitation given by the available data. It uses information on the distribution of Venezuelan immigrants by city/region of origin in a year after the shock initially started (2018). However, the RAMV is the only available source of information that I know contains the region of origin of Venezuelan immigrants. Therefore, I assume that the percentage of immigrants from each sending region in Venezuela has remained stable. This could be plausible considering the long history shared by the two countries.

⁶² The location of Colombians in Venezuela before the shock also serves as an information mechanism that conveys information about where to locate in Colombia.

⁶³ Indeed, Delgado-Prieto (2022), using a similar instrument, shows that the Rotemberg weights associated with the IV estimate are concentrated among two cities of origin in Venezuela: Maracaibo and San Cristobal. These two cities are indeed closer to other Colombian cities than other cities of origin. Maracaibo is close to states in the northern

closer to the border could have been more affected by economic shocks arising from changes in trade patterns between Colombia and Venezuela which could be correlated with changes in the economic outcomes of natives and therefore affect their location decisions.⁶⁴

Following equations (3) and (4) along with instruments given by equations (5) and (6), our empirical strategy seeks to determine whether differential exposure to rates of immigration from Venezuela affects working-age natives' location decisions. By focusing on changes, I can control for unobserved level differences. As long as there are no time-varying omitted local characteristics that correlate with our instrument and the outcome variable, our empirical strategy is thus valid.

To test the validity of the enclave instrument, I follow Goldsmith-Pinkham *et al.* (2020) recommendations.⁶⁵ First, I look at the correlation between the exposure shares given by $M_{c,2005}/M_{2005}$ in equation (5) and local characteristics in two years prior to the immigration shock (2005 and 2013). Table 2.2 presents the results of this exercise.

part of Colombia (La Guajira, Magdalena, and Cesar). In contrast, San Cristobal is close to the terrestrial entry at the City of Cucuta in Norte de Santander. Therefore, the instrument compares states in Colombia that received many immigrants from San Cristobal and Maracaibo with those that received fewer immigrants from these two cities in Venezuela.

⁶⁴ This concern might not be as important considering that shocks to trade patterns between Venezuela in Colombia occurred way before any immigration shock happened, which was caused by push factors in Venezuela. In particular, changes in economic and political conditions.

⁶⁵ In section 5.2, I use the distance instrument to check the robustness of our main estimates using the enclave instrument.

Table 2.2 Relationship between exposure shares and local characteristics in 2005 and 2013

Characteristics	Exposure share			
	Coef.	SE	Coef.	SE
Panel A. Local characteristics in 2005				
Share of population age 18 to 64	0.779	(0.612)		
Share of college-educated population	0.008	(0.374)		
Share of workers self-employed	-0.303	(0.442)		
Share of employment in manufacturing	-0.848	(0.629)		
Panel B. Local characteristics in 2013				
Natives mean age			0.247	(0.371)
Natives mean experience			-0.259	(0.372)
Natives mean schooling			-0.117	(0.116)
Natives fraction high skill			0.092	(1.487)
Natives share in manufacturing			0.415	(0.444)
Natives share in commerce			0.709	(0.498)
Natives share in real state			1.571	(1.068)
Natives share informal			0.057	(0.251)
Natives mean log hourly labor income			0.022	(0.085)
Obs.	23		23	
$\overline{R^2}$	0.237		0.723	

Notes: Each column reports the results of a single regression of the 2005 Venezuelan share on a set of local characteristics in 2005 and 2013. Results are weighted by the 2013 working-age population. Standard errors are in parentheses. *Source:* Own calculations using data from the GEIH.

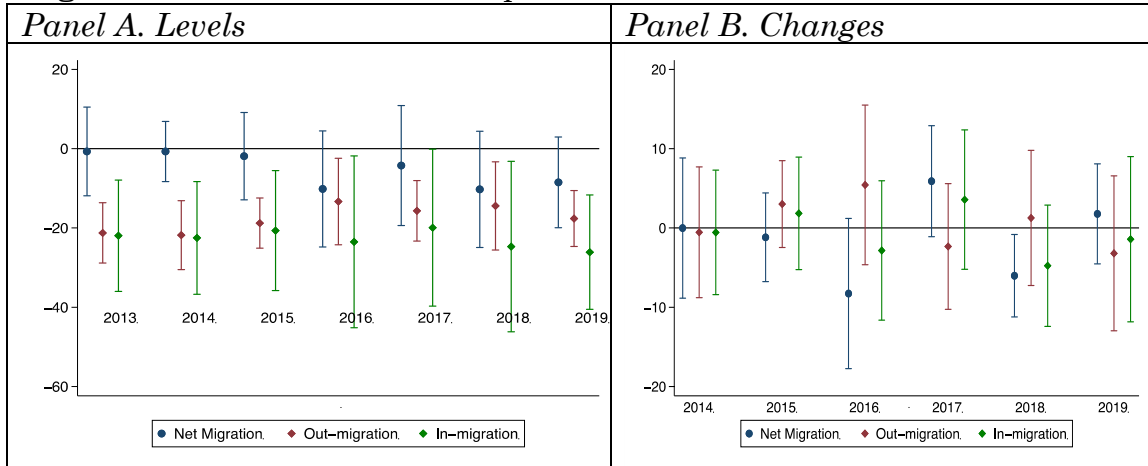
Local characteristics explain a good amount of the variability of the shares (24% in 2005 and 72% in 2013). However, there is no statistically significant relationship between shares and state characteristics in 2005 and 2013. Therefore, results presented in Table 2.2 suggest no other channels through which the shares affect outcomes in the relevant period except by predicting where the new inflows of immigrants locate within the country.

Second, I also test for pre-trends. The idea is to test whether the exposure shares can predict our outcomes of interest in the years before the immigration shock. In particular, whether exposure shares are related to levels or changes in net migration rates (and its different components, the in-migration and the out-migration rates) in the year before 2016 when the border between the two countries was reopened, and an exponential increase in the inflows occurred (see Figure 2.4 in the appendix). Figure 2.3 presents the results of this exercise. The figure reports year-by-year estimates of a reduced-form regression of migration rates (net migration rate, and its components, the in-migration and out-migration rates) on the share of Venezuelan immigrants in 2005. Panel A presents the results in levels of the outcomes, whereas Panel B does the same for changes (first differences).

Figure 2.3 shows that although the variation in the share of Venezuelan immigrants in 2005 did predict in and out-migration rates in levels, it did not once the equations were estimated in the first differences. Since the shares are equilibrium values, it is not uncommon that they are correlated with the levels of the outcomes. The critical issue for identification is that they are not connected to changes in the outcomes in years prior to the shock. Panel B in Figure 2.3 shows that, indeed, shares only predict changes in outcomes after 2015, suggesting that an important

change in the nature of the immigration inflows took place, potentially affecting natives' internal migration decisions.

Figure 2.3 Pre-trends for the exposure shares



Notes: The figures report year-by-year estimates of a reduced-form regression of migration rates (net migration rate, and its components, the in-migration and out-migration rates) on the share of Venezuelan immigrants in 2005. I present pre-trend results for the dependent variables in levels (Panel A) and first differences (Panel B). The sample size for every year is the number of states (23). Regressions include the set of controls described in Table 2.2, panel B. Estimates weighted by the working-age population in 2013. *Source:* Own calculations using data from the GEIH.

2.5 Results

2.5.1 Effects on population changes and migration rates

Following equation (3), I begin by presenting the effect of changes in the inflow of immigrants in consecutive years on the changes in the native population in Table 2.3. By construction, all the changes in a state's population between two consecutive years are explained by natives arriving or leaving the state and new immigrants from Venezuela moving to that state. Columns 1–3 present estimates using an OLS estimator, whereas columns 4–6 use a 2SLS estimator, instrumenting $\frac{F_{c,t}-F_{c,t-1}}{L_{c,2013}}$ in

equation (3) with the enclave instrument given by equation (5). Columns 2 and 5 include year and state-fixed effects. In contrast, columns 3 and 6 include the same controls as 2 and 5, but also time-varying state characteristics capturing demographic and economic conditions of a given state that could also affect changes in its population levels. The controls at the state level are the mean natives' age and potential experience and their squares, the fraction of natives with a college degree or more, the fraction of natives in the manufacturing sector, and the fraction of natives in the informal sector as defined by DANE.⁶⁶

At the bottom of Table 2.3, I report the strength of the first stage as a measure of the instrument relevance through the effective F-statistics proposed by Montiel and Pflueger (2013) that is robust to heteroscedasticity, serial correlation, and clustering. The latter is a common feature of panel data and, therefore, an important consideration in our empirical estimation. Although all the F-statistics are well above the conventional rule of thumb of 10 (Staiger & Stock, 1997), implying that weak instruments are less of a concern in our identification, as suggested by Andrews et al. (2019), I present identification-robust Anderson-Rubin (AR) confidence intervals for our point estimates that are efficient regardless of the strength of the first stage. Those confidence intervals are

⁶⁶ In section 5.2, I run additional specification to test the robustness of our main results to the inclusion of other variables such as state's GDP and population size.

reported in brackets below the standard errors of the point estimates in Tables 2.3 and 2.4.

Overall, results from Table 2.3 suggest that changes in the inflow of immigrants from Venezuela are negatively associated with changes in the native population across Colombian states. This is the case using either the OLS or 2SLS point estimates. Coefficients get more negative once the changes in Venezuelan immigration inflows are instrumented with the enclave instrument, highlighting the potential presence of measurement error in the immigrant counts and its implications for attenuation bias when using the OLS specification. However, the fact that coefficients did not change signs and are relatively similar could also indicate no evidence of endogeneity bias due to local demand shocks that drew immigrants to certain states.

Table 2.3 Effect of immigration from Venezuela on native population changes across states, 2013-2019

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Effect on natives' population changes</i>						
Change in immigrant population	-0.22	-0.26	-0.12	-0.46	-0.44	-0.32
	(0.14)	(0.13)	(0.07)	(0.19)	(0.18)	(0.17)
Andrews et al. (2019) AR 95% CI				[-0.79, -0.03]	[-0.73, 0.01]	[-0.66, 0.06]
Year FE	✓	✓	✓	✓	✓	✓
State FE		✓	✓		✓	✓
State Controls			✓			✓
Montiel & Pflueger (2013) Effective F-stat				29.7	24.9	19.0
N	138	138	138	138	138	138

Notes: The table reports the coefficients obtained by regressing year-to-year changes in native population on the change in immigrant inflows across states between 2013 and 2019 as defined in

equation 3. Models include year and state-fixed effects and a set of time-varying controls that also affect population changes. The controls at the state level are the mean natives' age and potential experience and their squares, the fraction of natives with a college degree or more, the fraction of natives in the manufacturing sector, and the fraction of natives in the informal sector as defined by DANE. Controls are similar to those included by Card (2001). I report 95% Anderson–Rubin confidence sets robust to weak instruments, as suggested by Andrews et al. (2019). The confidence sets were obtained using the package *twostepweakiv* in Stata 17 (Sun, 2018). At the bottom of the table, I report a measure of the quality of the first stage through the effective F-stat suggested by Montiel & Pflueger (2013). The effective F-stats were obtained using the *weakivtest* package in Stata 17 (Pflueger & Wang, 2013). Estimates are weighted by the working-age population in 2013. Cluster robust standard errors are reported in parentheses.

OLS and 2SLS show a negative statistically significant effect at 5% significance levels in specifications only using year and state-fixed effects. Our preferred specification given by column 6, which includes time-varying state controls, shows a negative and significant effect at a 10% level using standard inference procedures. However, once confidence sets robust to the presence of weak instruments are constructed (in brackets), the effect is not distinguishable from zero. Results from column 6 stress the importance of controlling for demographic and economic conditions in the states when explaining changes in native population between consecutive years and, more importantly, to report identification-robust Anderson-Rubin (AR) confidence intervals for the parameter of interest regardless of the value of the first stage as recommended by Andrews et al. 2019 (p. 729).

It is also important to highlight that our coefficient of interest seems less sensitive to the inclusion of time-varying controls, which could potentially be deemed “bad controls,” such as the fraction of natives with a

college degree which could also be affected by the immigration shock.⁶⁷ Therefore, although including controls measuring educational outcomes (such as schooling or the fraction with a college degree or more) could potentially induce a spurious correlation between changes in immigration and changes in native state populations, the stability in the coefficient of interest (in both sign and size), when controlling for this variable provides some reassurance that, at least in our context, this source of bias is less relevant and allows small gains in precision. Therefore, when discussing the results, my preferred specification controls for time-varying state characteristics.

Although aggregated changes in immigration from Venezuela might not cause overall changes in the net-native population across states, it is important to distinguish what drives this relationship. In Table 2.4, I present the effects of changes in immigration inflows on the different components of net migration rates, e.g., in-migration and out-migration rates. Panel A shows the results for natives' out-migration and Panel B for natives' in-migration rates, respectively. Changes in immigration inflows from Venezuela are positively associated with natives' out-migration and negatively associated with natives' in-migration rates in models controlling for state and year-fixed effects. Our preferred specification in column 6,

⁶⁷ Several literature connect immigration shocks to changes in natives' schooling decisions. For the particular case of Colombia, Rozo & Vargas (2020), show that the migration shock from Venezuela decreased promotion rates and increased dropout rates for both native and foreign students.

including state controls, shows that the effect on out-migration rates is statistically significant at 10% significance using standard inference.

Table 2.4 Effect of immigration from Venezuela on natives' in and out-migration rates across states, 2013-2019

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Effect on out-migration rates</i>						
Change in immigrant population	-0.46 (0.16)	0.13 (0.069)	0.12 (0.067)	-0.78 (0.24)	0.13 (0.12)	0.22 (0.12)
Andrews et al. (2019) AR 95% CI				[-1.17, -0.16]	[-0.22, 0.33]	[-0.07, 0.46]
Year FE	✓	✓	✓	✓	✓	✓
State FE		✓	✓		✓	✓
State Controls			✓			✓
<i>Panel B: Effect on in-migration rates</i>						
Change in immigrant population	-0.68 (0.23)	-0.13 (0.11)	-0.008 (0.07)	-1.24 (0.39)	-0.31 (0.12)	-0.08 (0.10)
Andrews et al. (2019) AR 95% CI				[-1.87, -0.25]	[-0.58, -0.04]	[-0.32, 0.11]
Year FE	✓	✓	✓	✓	✓	✓
State FE		✓	✓		✓	✓
State Controls			✓			✓
Montiel & Pflueger (2013) Effective F-stat				29.7	24.9	19.0
N	138	138	138	138	138	138

Notes: The table reports the coefficients obtained by regressing natives' out-migration (Panel A) and in-migration rates (Panel B) on the change in immigrant inflows across states between 2013 and 2019 as defined in equation 3. Models include year and state-fixed effects and a set of time-varying controls that also affect population changes. The controls at the state level are the mean natives' age and potential experience and their squares, the fraction of natives with a college degree or more, the fraction of natives in the manufacturing sector, and the fraction of natives in the informal sector as defined by DANE. Controls are similar to those included by Card (2001). I report 95% Anderson-Rubin confidence sets robust to weak instruments, as suggested by Andrews et al. (2019). The confidence sets were obtained using the package *twostepweakiv* in Stata 17 (Sun, 2018). At the bottom of the table, I report a measure of the quality of the first stage through the effective F-stat suggested by Montiel & Pflueger (2013). The effective F-stats were obtained using the *weakivtest* package in Stata 17 (Pflueger & Wang, 2013). Estimates are weighted by the working-age population in 2013. Cluster robust standard errors are reported in parentheses.

However, effects for both natives' in and out-migration rates are indistinguishable from zero once confidence sets robust to the presence of

weak instruments are constructed. Results imply that although the direction of the effects seems consistent with the idea that natives are inclined to leave states receiving high inflows of immigrants from Venezuela or to avoid migration into those states, I am not able to find a statistically significant relationship in our sample of states for the period 2014-2019.

How do our estimates compare to the international literature on the effects of migration inflows on natives' population changes? Our results align with the previous literature that uses similar empirical specifications for developed countries. In particular, our results align with Card (2001) and Card (2007), who show that the mobility flows of US natives are not that sensitive to the arrival of immigrants. Although other strands of literature show evidence of displacement in the US (see Borjas, 2006; Ali et al., 2012), Canada (Baine & Coulombe, 2018), and Italy in the case of low-skilled natives (Basile et al., 2021), empirical specifications in those documents are not directly comparable to ours, and some of them have important shortcomings. Indeed, Peri & Sparber (2011) show through simulations that specifications such as those implemented by Borjas (2006) could mechanically force a negative correlation between immigrant inflows and native population changes even when no displacement exists. On the other hand, specifications that do not normalize native and foreign-born population changes by population levels in a base year could generate a

spurious positive correlation between changes in immigration and changes in the native population. This seems to be the case with the empirical specifications defined by Ali et al. (2012), Baine and Coulombe (2018), and Basile et al. (2021). Something similar could be said for specifications that connect in and out-migration to immigration from Venezuela in Colombia, as discussed later.

My results also stand in contrast to Del Carpio et al. (2015), one of the few documents I am aware of studying native location responses to immigration in the developing world (Malaysia). They found that immigration induced positive inflows of natives. For every ten immigrants into a given state, there was an additional inflow of almost eight working-age natives, with most of the effect driven by natives finding new job opportunities and less by natives being out of the labor force. Despite similarities in the immigration profiles (at least in education and migratory status) between immigrants arriving in Malaysia and Venezuelan arriving in Colombia, I cannot detect native location responses for 2014-2019. There are several reasons why results for two developing countries experiencing immigration shocks differ. First, Del Carpio et al. (2015) also explore margins of adjustment outside the working-age population. They explore how the native labor supply responds to immigration by continuing education or retiring. I focus on working-age

natives' location responses in response to changes in the intensity of immigration.

Second, natives could respond to immigration shocks over a longer horizon. Therefore, the two decades of data used for the Malaysian case could have facilitated to capture natives' responses, in contrast to the Colombian case, where I only used variation between 2014-2019. Third, results could have also differed due to the specific differences in the labor markets receiving each shock, specifically regarding housing prices. Albert & Monras (2020) show that internal migration is more responsive in high than low housing price locations. There may have been no substantial differences in housing prices between cities attracting immigrants in Malaysia, which could have been the case in Colombia. However, in the Colombian case, it is challenging to connect housing prices to changes in natives' location response because the indexes measuring changes in housing prices in Colombia are built for 13 main metropolitan areas and not for states which are our unit of analysis.

How do the results relate to preliminary evidence of internal migration response to immigration from Venezuela? Several recent articles explore the role of internal migration as an adjustment mechanism in response to immigration from Venezuela in the context of providing robustness to the estimates of the effect of immigration on the wages of

low-skilled Colombians and other labor market outcomes (Bonilla et al., 2023; Lebow, 2022; Delgado-Prieto, 2022; Caruso et al., 2019). If natives respond rapidly to the arrival of Venezuelans to a given area, then estimates that do not account for those movements could underestimate the effect of immigration on natives' wages and other labor market outcomes. The evidence on the impacts of Venezuelan immigration on natives' location decisions is mixed. Bonilla et al. (2023) show that immigration reduces natives' in-migration, whereas Caruso et al. (2019) and Lebow (2022) find no statistically significant effects on either in-migration or out-migration.⁶⁸ However, as discussed in the introduction, their empirical specifications and variable constructions are limited and do not allow them to appropriately capture natives' location response to one additional immigrant arriving at a given location.

I improve over previous preliminary evidence by connecting natives' location responses to immigration from Venezuela through empirical specifications well designed to identify native location responses to

⁶⁸ Bonilla et al. (2023) run models for the probability of in or out-migration and show that immigration reduces the probability of in-migrating, implying that immigrants crowd out natives from areas with high immigration. Caruso et al. (2019) ran LPM for in-migration and out-migration. They found a negative correlation between immigration shares (defined as the fraction of Venezuelan immigrants in the working-age population) and natives' in-migration and positively correlated to immigrants. However, the effects were not statistically significant. Lebow (2022) ran a first-difference specification between 2014-2019 relating the effect of Venezuelan arrivals on the normalized count of natives who entered (in migration) or left (out-migration) a metro area over the previous five years as a share of the 5-year lagged population. He found similar results to Caruso et al. (2019).

immigration (as discussed by Peri & Sparber, 2011) and by also exploiting year-to-year variation across geographic units as been standard practice in most of the documents interested in the labor market effects of Venezuelan immigration in Colombia. Contrary to Bonilla et al. (2023), but in line with Caruso et al. (2019) and Lebow (2022), I find no nonexistent offsetting native responses to Venezuelan immigration over the period 2014-2019. Results are important to confirm that previous findings on the effects of Venezuelan immigration on natives' wages, especially for those less educated (Caruso et al., 2019; Lebow, 2022), were not biased due to natives' moving out of or avoiding places with a higher concentration of immigrants from Venezuela.

2.5.2 Robustness Checks

I conduct additional exercises using alternative regression specifications and different instruments to assess the robustness of estimates presented in Tables 2.3 and 2.4. Table 2.5 reports the results of the robustness exercises. I report all coefficients with their corresponding standard errors, the first stage effective F-statistics proposed by Montiel & Pflueger (2013), and identification-robust Anderson-Rubin (AR) confidence intervals for the parameter of interest as suggested by Andrews et al. 2019.

Table 2.5 Robustness exercises for the effect of changes in immigration from Venezuela on changes in natives' net migration rates and its components

<i>Estimates of the effect of changes in immigrant population on:</i>	Coef.	SE	<i>Montiel & Pflueger (2013) Effective F-stat</i>	Andrews et al. (2019) AR 95% CI
<i>(1) Changes in native population growth</i>				
(a) Using a distance-based instrument	-0.06	0.27	11.0	[-0.77, 0.56]
(b) Using the shares from the 1993 census	-0.26	0.12	14.1	[-0.50, 0.04]
(c) Model including additional controls	-0.55	0.23	15.1	[-1.12, -0.06]
(d) Normalizing variables by t-1 working age population	-0.30	0.17	17.9	[-0.67, 0.08]
<i>(2) Native out-migration rates</i>				
(a) Using a distance-based instrument	-0.14	0.14	11.0	[-0.65, 0.11]
(b) Using the shares from the 1993 census	0.17	0.09	14.1	[-0.06, 0.33]
(c) Model including additional controls	0.30	0.15	15.1	[-0.01, 0.69]
(d) Normalizing variables by t-1 working age population	0.23	0.14	17.9	[-0.09, 0.51]
<i>(3) Native in-migration rates</i>				
(a) Using a distance-based instrument	-0.20	0.21	11.0	[-0.94, 0.18]
(b) Using the shares from the 1993 census	-0.08	0.95	14.1	[-0.29, 0.10]
(c) Model including additional controls	-0.25	0.16	15.1	[-0.62, 0.15]
(d) Normalizing variables by t-1 working age population	-0.07	0.11	17.9	[-0.33, 0.17]

Notes: The table reports various estimates of the effect of changes in immigration inflows on natives' changes in population and migration rates across states between 2013 and 2019. I report 95% Anderson–Rubin confidence sets robust to weak instruments, as suggested by Andrews et al. (2019), and a measure of the quality of the first stage through the effective F-stat as suggested by Montiel & Pflueger (2013). N = 138 for all the regressions estimated. Estimates are weighted by the working-age population in 2013. Cluster robust standard errors are reported beside the point estimate.

In rows (a) and (b), I report estimates for our preferred specification (column 6 in Tables 2.3 and 2.4) using two different instruments: a distance-based instrument as defined in equation 6, and an enclave instrument as the one presented in equation 5 but built using exposure

shares coming from the 1993 census. In row (c), I include additional controls, such as the state GDP and the log of the working-age population. In row (d), I normalize the outcome variables and the changes in immigration from Venezuela by the state's working-age population in $t - 1$.

Taken together, results from Table 2.5 suggest, as presented in Tables 2.3 and 2.4, that changes in immigration from Venezuela do not significantly affect natives' internal migration patterns. Results from Panel (1) suggest that although the point estimates on net migration rates are consistently negative across specifications and potentially statistically significant at the 5% level (using the enclave instrument based on the shares of the 1993 census) by traditional standard inference, the effect is not distinguishable from zero once identification robust confidence intervals are constructed. However, one specification, row (c) that includes state GDP and the log of the working-age population as additional controls, shows evidence of displacement in terms of net migration rates but fails to identify whether the effect is explained by changes in in-migration or out-migration rates (see row (c) on Panels (2) and (3)). Specification including these controls could be problematic because both could be affected by the shock and, therefore, be deemed "bad controls." For this reason, I do not include them in our main specifications discussed in Tables 2.3 and 2.4.

2.6 Concluding remarks

Even though more than 35% of the migration flows happen between developing countries, most of the literature on the impact of international migration supply shocks on internal mobility has focused on how developed countries absorb these shocks. Results have shown that native responses go from moderately or nonexistent offsetting responses to international immigration in the US and France to displacement effects in the cases of Germany, Italy, and Canada. To the extent to which natives in developing countries could react similarly to previous experiences and under what conditions this happened is still an open question.

In this document, I provide evidence for the developing world. I study one of the most prominent south-to-south migration episodes in the last 20 years: the arrival of immigration from Venezuela in Colombia over the period 2014-2019. Using plausible exogenous variation in the arrival of Venezuelans in Colombia and data on natives' location responses, I found no evidence that immigration from Venezuela induces changes in net-migration rates across Colombian states. In short, on net, the arrival of immigrants into states does not induce population loss or gain across states over 2014-2019. I also found evidence that although changes in immigration rates are positively associated with natives' out-migration and negatively associated with natives' in-migration rates, the effects are

not distinguishable from zero once confidence sets robust to weak instruments are constructed. Highlight the importance of presenting robust inference when using IV approaches, as Andrews et al. (2019) discussed.

There are limitations of our analysis that are important to highlight. First, there could be important differences in native and Venezuelan immigrant inflows happening even inside states. If local labor market responses at the population level are happening between municipalities within states and not between states, then it could be possible that our overall conclusions regarding the role of immigration on natives' mobility rates and, therefore, its implications for labor market outcomes analysis are overly optimistic, meaning that either displacement (increases in out-migration) or crowding out effects (decreases in in-migration) could still be at play. Second, it could also be possible that native migration patterns are insensitive to immigration from Venezuela in the 12-month time frame implied by our analysis and that any native mobility response is more long lagged.

Third, there could be heterogeneity in the native responses depending on the characteristics of the states. For instance, Albert and Monras (2020) show that native mobility responses to immigration are higher in areas with higher housing prices. Considering that immigrants

from Venezuela are, on absolute terms, concentrated in the most populated metropolitan areas with higher rent and housing prices, this could be a future avenue of research to pursue. At the moment, however, data on rent and housing prices are only representative of 13 metropolitan areas and, therefore, incompatible with our specifications at the state level. Finally, my findings could also be validated or complemented by research using variation between the 2005 and 2018 population censuses, which could be closer in spirit to influential papers in the literature on natives' location responses to immigration shocks (Card, 2001; Card, 2007).

However, our analysis is still important to understand the short-term responses to immigration supply shocks in the developing world and for the overall literature on the labor market effects of immigration from Venezuela across Colombian local labor markets since most of the empirical specifications in this area use variation across years within local labor market defined by MSAs or states. Our findings offer some validity that previous research showing that negative effects of immigration on low-skilled natives' wages are not biased by natives' internal mobility responses.

REFERENCES

- Abramitzky, R., Ager, P., Hansen, C.W., Cohen, E., & Boustan, L.P. (2021). The Effect of Immigration on Local Labor Markets: Lessons from the 1920s Border Closure. KC FED Research Working Papers, RWP 21-09.
- Albert, C., & Monras, J. (2020). The Regional Impact of Economic Shocks: Why Immigration is Different from Import Competition. Economics Working Papers 1758, Department of Economics and Business, Universitat Pompeu Fabra.
- Ali, K., Partridge, M.D., & Rickman, D.S. (2012). International immigration and domestic out-migrants: are domestic migrants moving to new jobs or away from immigrants? *The Annals of Regional Science*, 49, 397–415.
- Altonji, J., & Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In J. M. Abowd & R. B. Freeman (Eds.), *Immigration, Trade, and the Labor Market* (pp. 201–234). National Bureau of Economic Research.
- Amior, M. (2021). Immigration, Local Crowd-Out and Undercoverage Bias. CEP Discussion Paper No 1669.
- Andrews, I., Stock, J.H., & Sun, L. (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics*, 11, 727–753.
- Aydemir, A., & Borjas, G. J. (2011). Attenuation bias in measuring the wage impact of immigration. *Journal of Labor Economics*, 29(1), 69–112.
- Barón, J. (2011). Sensibilidad de la oferta de migrantes internos a las condiciones del mercado laboral en las principales ciudades de Colombia. *Documentos de Trabajo Sobre Economía Regional y Urbana*, 149.
- Basile, R., De Benedictis, L., Durbán, M., Faggian, A., & Mínguez, R. (2020). The impact of immigration on the internal mobility of natives and foreign-born residents: evidence from Italy. *Spatial Economic Analysis*, 16, 9–26.
- Beine, M., & Coulombe, S. (2018). Immigration and internal mobility in Canada. *Journal of Population Economics*, 31, 69–106.

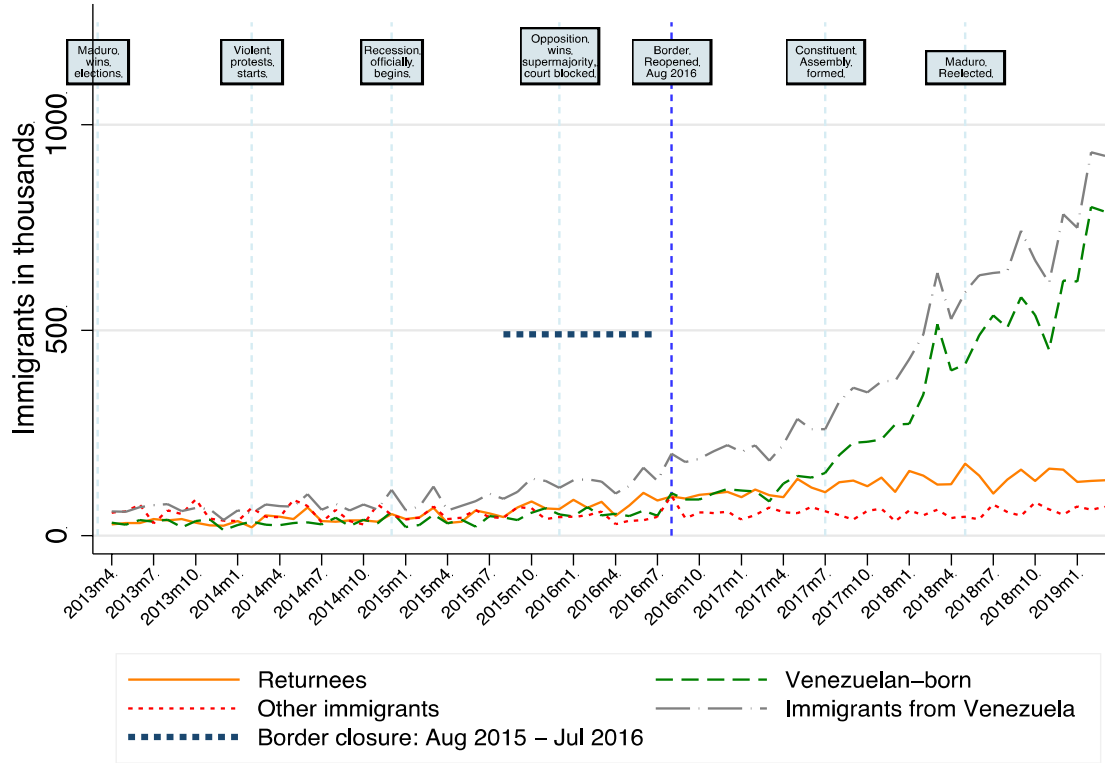
- Bonilla-Mejía, L., Morales, L.F., Hermida, D., & Flórez, L.A. (2023). The Labor Market Effect of South-to-South Migration: Evidence from the Venezuelan Crisis. *International Migration Review*, 1–36.
- Borjas, G. J. (2006). Native Internal Migration and the Labor Market Impact of Immigration. *The Journal of Human Resources*, XLI, 221–258.
- Borjas, G. J., Freeman, R. B., Katz, L. F., DiNardo, J., & Abowd, J. M. (1997). How Much Do Immigration and Trade Affect Labor Market Outcomes? *Brookings Papers on Economic Activity*, 1997(1), 1.
- Boustan, L. P., Fishback, P. V., & Kantor, S. (2010). The Effect of Internal Migration on Local Labor Markets: American Cities during the Great Depression. *Journal of Labor Economics*, 28(4), 719–746.
- Bozzoli, C.G., Brück, T., & Wald, N. (2011). Self-employment and Conflict in Colombia. *Journal of Conflict Resolution*, 57, 117–142.
- Burstein, A.T., Hanson, G.H., Tian, L., & Vogel, J. (2020). Tradability and the Labor-Market Impact of Immigration: Theory and Evidence from the United States. *Econometrica*, 88, 1071–1112.
- Calderón-Mejía, V., & A. M. Ibáñez. (2016). Labour Market Effects of Migration-Related Supply Shocks: Evidence from Internal Refugees in Colombia. *Journal of Economic Geography*, 16(3), 695–713.
- Card, D. (2001). Immigrant Inflows, Native Out-flows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, 19(1), 22–64.
- Card, D. (2007). How Immigration Affects US Cities. Working Paper, Centre for Research and Analysis of Migration. London, UK.
- Card, D., & DiNardo, J. (2000). Do Immigrant Inflows Lead to Native Out-flows? *American Economic Review Papers and Proceedings*, 90(2), 360–367.
- Caruso, G., Gomez-Canon, C., & Mueller, V. (2019). Spillover Effects of the Venezuelan Crisis: Migration Impacts in Colombia. *Oxford Economic Papers*, 1–25.
- Colas, M. (2018). Dynamic Responses to Immigration. FRB Minneapolis, OIGI Working Paper No. 6.

- Cortes, P. (2008). The Effect of Low-Skilled Immigration on US Prices: Evidence from CPI Data. *Journal of Political Economy*, 116(3), 381–422.
- Del Carpio, X., & Wagner, M. (2015). The Impact of Syrian Refugees on the Turkish Labor Market. World Bank Policy Research Working Paper No. 7402, Available at SSRN: <https://ssrn.com/abstract=2650218>
- Del Carpio, X., Özden, Ç., Testaverde, M., & Wagner, M. (2015). Local Labor Supply Responses to Immigration. *The Scandinavian Journal of Economics*, 117(2), 493–521.
- Delgado-Prieto, L (2022). Immigration, Wages, and Employment under Informal Labor Markets. UC3M Working papers, Economics 35664, Universidad Carlos III de Madrid, <https://e-archivo.uc3m.es/bitstream/handle/10016/35664/we2212.pdf?sequence=1>.
- Dustmann, C., Tuhler, J., & Schönberg, U. (2017). Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment. *Quarterly Journal of Economics*, (September), 435–483.
- Edo, A. (2020). The Impact of Immigration on Wage Dynamics: Evidence from the Algerian Independence War. *Journal of the European Economic Association*, 18(6), 3210–3260.
- Filer, R. (1992). The Effect of Immigrant Arrivals on Migratory Patterns of Native Workers. In G. J. Borjas & R. B. Freeman (Eds.), *Immigration and the Workforce: Economic Consequences for the United States and Source*. Chicago University Press.
- Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8), 2586–2624.
- Hatton, T.J., & Tani, M. (2005). Immigration and Inter-Regional Mobility in the UK, 1982-2000. *Economic Journal*, 115(507), 342–358.
- Lebow, J. (2022). Immigration and Occupational Downgrading in Colombia. Working Paper, <https://static1.squarespace.com/static/6064ff1093aaa5277d9d8df9/t/6370129dd35f087529990bbf/1668289188429/Lebow2021.pdf>

- Monras, J. (2020). Immigration and Wage Dynamics- Evidence from the Mexican Peso Crisis. *Journal of Political Economy*, 128(8), 3017–3089.
- Monras, J. (2021). Local Adjustment to Immigrant-Driven Labor Supply Shocks. *Journal of Human Capital*, 15, 204–235.
- Morales, J. (2017). The Impact of Internal Displacement on Destination Communities: Evidence from the Colombian Conflict. *Journal of Development Economics*, 131, 132–150.
- OECD (2016). *International Migration Outlook 2016*, OECD Publishing, Paris. http://dx.doi.org/10.1787/migr_outlook-2016-en
- Olea, J.L., & Pflueger, C.E. (2013). A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics*, 31, 358–369.
- Peri, G., & Sparber, C. (2011). Assessing Inherent Model Bias: An Application to Native Displacement in Response to Immigration. *Journal of Urban Economics*, 69(1), 82–91.
- Pflueger, C.E., & Wang, S. (2013). WEAKIVTEST: Stata module to perform weak instrument test for a single endogenous regressor in TSLS and LIML. *Statistical Software Components*.
- Rozo, S.V., & Vargas, J.F. (2020). The Effect of Venezuelan Migration on Educational Outcomes in Colombia. IDB-DP-80.
- Sun, L. (2018). Implementing Valid Two-Step Identification-Robust Confidence Sets for Linear Instrumental-Variables Models. *The Stata Journal*, 18, 803 - 825.
- Weber, S., & Péclat, M. (2017). A Simple Command to Calculate Travel Distance and Travel Time. *The Stata Journal*, 17(4), 962–971.
- World Bank (2018). *Migración desde Venezuela a Colombia: Impactos y Estrategias de Respuesta en el Corto y Mediano Plazo*. The World Bank Group, October 2018.

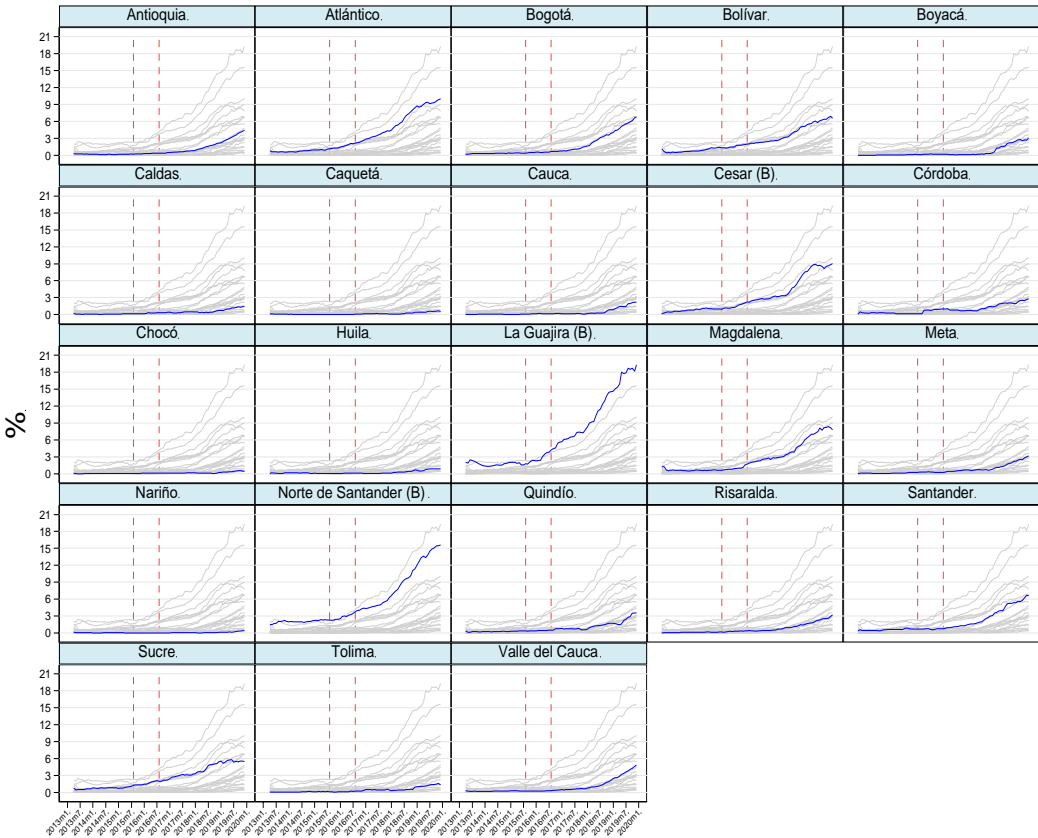
APPENDIX

Figure 2.4 Monthly stock of immigrants from Venezuela across Colombian MSAs



Notes: The Figure depicts the evolution of the monthly stock (in thousands) of working-age (15-64 years old) immigrants from Venezuela and other immigrant groups living in urban areas between April 2013 and March 2019. The gray dashed line corresponds to the sum of all migrants coming from Venezuela, including returnees and Venezuelan-born. The horizontal blue dashed line indicates when the Venezuelan government unilaterally decided to close the main border crossings between the two countries. The vertical dashed blue line indicates the date when the border was reopened. Light-blue boxes on the upper part of the graph show important political and economic events in Venezuela. Returnees are defined as Colombian-born individuals returning from Venezuela. *Source:* Own estimates using data from the GEIH 2013-2019.

Figure 2.5 Evolution in the share of immigration from Venezuela in the native working-age population 2013-3 – 2019-12



Notes: The figure depicts the monthly evolution of the fraction of immigrants from Venezuela in the native working-age population (18-64 years old) across Colombian states. The counts of natives and immigrants from Venezuela used to compute the fraction are 12-month moving averages. The counts were adjusted by population weights based on the 2005 Population Census. Immigrants from Venezuela are the sum of the Colombians returning to the country (returnees) and Venezuelan-born immigrants. The dashed red vertical lines in the graphs show the period when the terrestrial border between Colombia and Venezuela was closed (August 2015 to July 2016). *Source:* Own estimates using data from the GEIH 2013-2019.

CHAPTER 3

IS THE CCN-51 CACAO HYBRID THE SOLUTION FOR THE
POTENTIAL SHORTAGE OF CHOCOLATE? A HEDONIC AND
WILLINGNESS TO PAY STUDY FOR DIFFERENT VARIETIES OF
ECUADORIAN CHOCOLATE**3.1 Introduction**

Chocolate has been experiencing a significant increase in demand during the last decade. The chocolate industry had a retail market value of 106 billion in 2017 and it is expected to increase to 190 billion by 2026 (MarketWatch, 2019). Two factors have driven the industry's growth. First, the emergent popularity of chocolate in countries with big markets and increasing income per capita, such as India, China, and Russia. Second, the recent evidence regarding the health benefits of certain types of chocolate (Moramarco & Nemi, 2016).⁶⁹ However, the future sustainability of the industry is in jeopardy. The current powerful boost in demand and the expectations of an even higher one might not be accompanied by an equal increase in supply (Gilbert, 2016, p. 321).⁷⁰ Cacao, the main component of chocolate and the one that gives its texture and flavor, is

⁶⁹ For instance, dark chocolate is rich in antioxidants, which are alleged to prevent cardiac diseases. These perceived health benefits have been driving demand's growth for products with larger cocoa concentration.

⁷⁰ The global chocolate market was valued at USD 113 billion in 2021, and it's expected to increase at a compound annual growth of 3.7% from 2022 to 2030 (Gran View Research, 2021).

facing various threats from climate change to fungal diseases to political instability in producing regions that can lead to poor yield returns and disincentivize farmers working in the cocoa production (Schmitz & Shapiro, 2012, p. 61; Läderach et al., 2013; Phillips-Mora & Wilkinson, 2007).

Several solutions have been proposed to boost cocoa production in anticipation of the potential shortage of chocolate. Selective breeding and farmer education efforts have been among the alternatives.⁷¹ However, less attention has been placed on exploring the potential for promoting the production of chocolate using cacao hybrids that have shown good yield and disease resistance, such as the CCN-51 cacao hybrid. In this paper, we address this gap in the literature and provide experimental evidence of the potential for chocolate made from CCN-51 to substitute high-quality chocolate. In particular, in an experimental lab setting, we study the hedonic properties and the factors affecting the WTP for Ecuadorian chocolate produced with different cocoa varieties, including the CCN-51. We elicit consumer preferences using the Becker-DeGroot-Marschak mechanism (BDM) (Becker et al., 1964). BDM is an incentive-compatible method in which subjects face a real purchase situation and has been

⁷¹ Selective breeding is used to find where in the world the cacao crops are disease resistant, more productive, and adaptable to different climate conditions. The education efforts pretend to teach farmers about novel planting, irrigation, and pest-management techniques.

extensively used to elicit WTP for foods (Noussair et al., 2004; Lusk & Fox, 2003; Kass & Ruprecht, 2006).

The CCN-51 variety was developed in Ecuador by Homero Castro in the 1970s and is now recognized worldwide as one of the most productive cocoa varieties (Boza et al., 2014). It produces around 1 to 3 tons per hectare, depending on the management practices (Amores et al., 2011). The hybrid also has several other benefits. First, it is resistant to the Witch's Broom disease. Second, CCN-51 trees start producing after two years, compared to 3-5 years for other varieties such as Trinitario. Third, CCN-51 trees are relatively short, averaging 5 meters. Finally, CCN-51 contains 54% cocoa butter and many antioxidants, which are highly demanded features in the chocolate and pharmaceutical industries.

Notwithstanding all the advantages over other cocoa varieties, using CCN-51 to make chocolate faces several challenges. On the one hand, some of the more prominent chocolate manufacturers (e.g., Hershey, Mars, and Cadbury) and experts have characterized its taste as acid and unsuitable for chocolate.⁷² Although new fermentation techniques have improved the acceptance among chocolate companies and chocolate tasters, many still think that the CCN-51 variety could affect the quality of the chocolate and consider that even the best CCN-51 is yet average in quality (Schatzker,

⁷² <http://www.smithsonianmag.com/science-nature/how-save-chocolate-tree-without-sacrificing-flavor-180954148/>

2014). Indeed, some countries regarded as producers of high-quality cocoa (e.g., Ecuador and Colombia) do not have the incentives to promote its cultivation because their international reputation could be affected. Therefore, it will put at risk the local producers of high-quality cocoa.⁷³ Finally, since the CCN-51 is genetically modified cocoa, its use in chocolate must deal with the current political and cultural disavowal of genetically modified foods (GMF) (Greenpeace International, 2008; Rommens, 2010).⁷⁴

Obstacles to using and promoting the CCN-51 variety are based on the premise that if customers perceive the chocolate made with CCN-51 as low quality or associated with GMF, their willingness to pay for it will be lower compared with chocolate manufactured with higher quality cocoa. However, as far as the authors know, there is no literature about the acceptance and WTP for chocolate elaborated with certain hybrids of cacao that have shown good yield and disease resistance (e.g., CCN-51) but that are considered of less quality by manufacturers. We contribute to this literature with empirical evidence on the hedonic valuation and WTP for chocolate elaborated with CCN-51 cacao.

⁷³ For instance, the Royal Tropical Institute (KIT) shows that countries that mixed the fine flavor cocoa with certain varieties such as CCN-51 receive a negative assessment, thus reducing their chances of being considered high potential countries for fine flavor cocoa (van der Kooij, 2013).

⁷⁴ Some literature has established that consumers do not have a good understanding about the GMF (Anand et al., 2007; Curtis et al., 2004).

Providing evidence on how consumers perceive chocolate made with CCN-51 cacao and how the perception is influenced by information available to them, and the interaction between hedonic properties, information, and the observed WTP, it is important to determine whether Chocolate made with CCN-51 cacao could be considered as an alternative to mitigate the potential shortage of chocolate, which in turn might help to pivot the policy efforts intended by manufacturers, universities, and government agencies.

This paper also contributes to the literature on the WTP for Genetically Modified Chocolate (Dannenberg et al., 2008; Kajale & Becker, 2014; and Noussair et al., 2002). Chocolate made from CCN-51 is, by definition, a genetically modified food. This literature shows that WTP for non-GM foods is higher than WTP for GM foods. Still, the effect is conditioned on the consumers' capacity to identify GM labeling and the type of information they receive. Using an experimental auction and a random sample of the residents of Mannheim (Germany), Dannenberg et al. (2008) found that the mean bid difference between a GM chocolate bar and its non-GM version was 0.26 euros.

Kajale and Becker (2014) study the effect of information on consumers' WTP for GMF using a convenience sample of university students in Delhi (India). Applying the Vickrey second price experimental

auction method to elicit WTP, authors found that consumers were willing to pay the highest amount under a combination of positive and negative information about GMF for a soya-chocolate bar. On the other hand, only positive or negative information had no positive or negative effect on WTP for GMF. Noussair et al. (2002), using the Vickrey second price experimental auction method to elicit WTP for a representative sample of consumers in Grenoble (France), found that labeling chocolate bars as containing GMOs does not affect consumers' WTP. Nevertheless, when the label is actually noticed, it induces a 30% decrease in WTP.

Finally, our document is also related to the broader literature on the effect of reduced fat labeling content on consumers' liking and WTP and to the literature related to value-based labeling for organic and fair-trade labels. In most of this literature, chocolate bars are used as products during the experiments, but with no specific implications for the chocolate industry and how the results could contribute to potential solutions to the future supply shortage of chocolate. We add to this literature by providing results that are relevant to the chocolate industry.

Within the first group, regarding the impact of reduced fat labeling on consumer perception, Norton et al. (2013) for the case of Birmingham (UK) and Kahkonen et al. (1999) using university employees and students in Finland, found that despite the fact that the expected pleasantness of

chocolate bars is lower for those with a low-fat label, the actual hedonic ratings were unaffected by the labeling. Norton et al. (2013) also found that liking was positively associated with participants' buying intentions and the price they would be willing to pay.

On the other hand, Tagbata and Sirieux (2008) estimate the WTP for organic and fair-trade chocolate bars in Montpellier (France) using the Becker-DeGroot-Marschak's mechanism (BDM) to elicit WTP. According to these authors, consumers' actual WTP for organic and fair-trade chocolate is higher than for traditional chocolate. For many consumers, the valuation of the organic and fair-trade labels is determined by the product's taste. Using a Discrete Choice Experiment and a larger sample size, Poelmans and Rousseau (2016) reached a similar result for Belgian consumers but with a markedly higher WTP. The respondents were willing to pay a premium of 10.60 euros per 250 gm of fair-trade chocolate compared to an identical product without a fair-trade label.

The remainder of the paper is organized as follows: Section 2 provides an overview of the cacao and chocolate industries and the threats to their future supply. Section 3 explains the experimental design. Section 4 describes the data and the empirical strategy. Section 5 depicts the results of the empirical estimations. Section 6 concludes and discusses policy implications and limitations.

3.2 Background on the cocoa and chocolate markets

Cocoa is predominantly a smallholder crop; over 90% of world cocoa production originates from small farms (ICCO, 2013). However, the trading, processing, and manufacturing activities are dominated by a few multinationals (Fold & Neilson, 2016). There are marked differences between the places where the cocoa is produced and the places where the final chocolate product is processed and consumed. As of 2021, 75% of the cocoa beans are produced in four African countries (Côte d'Ivoire, Ghana, Cameroon, and Nigeria) and 15% in Latin American countries (primarily concentrated in Ecuador and Brazil) (ICCO, 2022).⁷⁵

Meanwhile, the processing industry is located near the consumption centers. 52% of the cocoa beans are processed and transformed into intermediate and final chocolate in the developed world (Western Europe and North America). However, during the last 20 years, the processing industry has expanded to countries previously related to only production tasks, such as Côte d'Ivoire, Malaysia, Ghana, Brazil, and Indonesia (Kox, 2000). In 2010, these countries accounted for almost 35% of the processing

⁷⁵ There are three general cocoa varieties: Criollo, Forastero, and Trinitario (<https://www.barry-callebaut.com/chocophilia/theobroma-cacao-food-gods>). The Criollo tree is originated from Central to South America, highly demanded for the finest chocolates, and represents only the 5% of the world's production. The Forastero tree is the most commonly grown cocoa, it is mainly cultivated in Africa, Ecuador, and Brazil, highly resistant to diseases and accounts for 80% of the world's production. Finally, the Trinitario is a natural hybrid created from cross-pollination that combines the durability of the Forastero and the refined taste of the Criollo.

industry.⁷⁶ This shift is partially explained by changes in the decomposition of the value added in the industry. Only 3.3% of the consumer spending on finished chocolate products goes to cocoa farmers. Of the remaining 96.7%, 22% goes to chocolate processing, and 26% goes to wholesale and retail. So, some producing countries have started making intermediate products and final chocolate to capture a more significant share of the industry's total value added.

Chocolate consumption has also reshaped during the last 20 years. During the period 2000-2013, chocolate confectionery consumption in kg/capita declined or stagnated in wealthier countries such as USA, Sweden, France, and Belgium but has grown in emerging countries such as Brazil, Russia, India, and China (Poelmans & Swinnen, 2016, p. 38).

Cacao production has been facing various threats across the world. First, in Côte d'Ivoire and Ghana, where 63% of the world's cocoa production is concentrated (ICCO, 2022), poverty and political instability are pushing the farmers out of the cocoa business. In Côte d'Ivoire, the per capita daily cocoa income in 2014 was \$1.17, significantly below the \$1.9 a day threshold used by the World Bank in 2015 to classify extremely poor people (Balineau, Bernath, and Pahuatini, 2016). Poverty means that farmers cannot afford fertilizers, pesticides, and fungicides that boost tree

⁷⁶ Calculations are based on Table 2.4 in Poelmans and Swinnen (2016, p. 27).

productivity. Even when they can afford it, they do not have the training to use it effectively. Hence, they sometimes opt to harvest a more profitable crop, such as rubber. The dire economic conditions are exacerbated by the fact that Côte d'Ivoire and Ghana, although supplying most of the cocoa beans necessary to elaborate chocolate products, only earn between 3% to 6% of the total value added of the industry.

Second, climate change has led to radical decreases in land suitable for cacao crops. For instance, in Côte d'Ivoire and Ghana (the first and second-world producers), some current cocoa areas could be unsuitable for production in the future due to dry weather (Läderach et al., 2013; Schroth et al., 2016). Indeed, some areas previously used for cocoa are now used for other crops (e.g., rubber). In Indonesia, the annual rains are becoming more intense over shorter periods, knocking the flowers off the cacao trees and avoiding pod formation (Schmitz & Shapiro, 2012, p. 61; Santosa et al., 2018).

Finally, several fungal diseases have been affecting the production of cacao across the world. For instance, estimates put the total cacao production loss at around 40% in 2012 (Ten Hoopen et al., 2012). In South America, it is common for farmers to lose production due to witches' broom and frosty pod rot.⁷⁷ The latter is the most recent and dangerous because it

⁷⁷ Brazil has known by first-hand the effect of one these fungal diseases in the production.

has been in the diffusion phase, probably due to increased human-mediated spread (Phillips-Mora & Wilkinson, 2007). Meanwhile, although not too recent, the former has played an essential role in the fall of production in countries such as Ecuador and Brazil. In West Africa and Asia, new pathogens, such as the swollen shoot virus and the ceratobasidium, are emerging, respectively.⁷⁸

Another critical element of the chocolate value chain is the emergence of the CCN-51 variety. Although initially discovered and produced in Ecuador in the 70s, production has recently shifted to other South American countries, such as Peru, Colombia, and Brazil. In some countries, the CCN-51 production is well established: in 2012, 32% of Ecuador's cocoa exports and 36% of Peru's exports were represented by this variety (Van der Kooij, 2013).

Despite scientific evidence that CCN-51 cacao is disease resistant and offers more yield (Boza et al., 2014; Amores et al., 2011), most supply countries (especially those in Latin America where weather conditions are favorable) have not made great efforts to increase its production. This could be explained by different reasons. On the one hand, some Latin American countries have been historically recognized as high-quality cocoa

During the 90s the witches' broom reduced the production by an estimated of 75 percent (<http://www.npr.org/templates/story/story.php?storyId=91479835>).

⁷⁸ For an extensive review of the characteristics of the different diseases and their relative impact across the different producing countries see Ploetz (2016).

producers (e.g., Ecuador and Colombia), thus mixing the national variety (i.e., the highest quality one) with the CCN-51 variety might affect their international reputation, and therefore, put in danger the local producers of high-quality cocoa.

3.3 Experimental Design

We ran an experiment with students at Cornell University. The subjects were recruited through the Lab for Experimental Economics and Decision Research (LEEDR) email system. The students participating received a compensation of \$10. The Becker-DeGroot-Marschak mechanism (BDM) (Becker et al., 1964) was used to elicit Willingness to Pay (WTP) for three varieties of Ecuadorian chocolate: one made with high-quality cocoa beans (National Ecuadorian Chocolate), another made from CCN-51 cocoa beans (CCN-51), and the last one made with banana and CCN-51 cocoa beans (CCN-51 Banana). Henceforth these three varieties will be called National, CCN-51, and CCN-51 Banana.⁷⁹

WTP for the chocolate varieties is collected from two sources of treatment. The students participating in the experiment (110 in total) were randomly assigned to one of the two group treatments, henceforth called the without information group (NO INFO) and with information

⁷⁹ BDM is an incentive-compatible method in which subjects face a real purchase situation and has been extensively used to elicit WTP for foods (Noussair et al., 2004; Lusk and Fox, 2003; Kass and Ruprecht, 2006).

group (INFO). The students in the first treatment did not receive any information about the three varieties of chocolate. They only revealed their WTP based on the smell and taste of the chocolate varieties. Appendix A1 presents the instructions given to the NO INFO group during the taste experiment. On the other hand, for the second treatment, the students were told attributes of the three chocolate varieties, such as the quality of the cocoa beans used in the manufacturing process and the amount of potassium and fat content, so they decided their WTP based on the hedonic properties and the information provided. Appendix A2 presents the instructions given to the INFO group during the taste experiment.

Subjects were seated randomly at individual computer terminals with privacy shields, were informed that all decisions they made would be kept strictly confidential and were paid \$10 on average to participate. 24 computers were available per session, and six sessions ranging from 15 to 24 subjects were implemented. After signing a consent form, participants were given a brief introduction to the experiment, which included the amount of money they would earn and the rules. Since previous work has shown that inexperienced subjects could submit bids that would not reflect their actual valuations (Kagel, 1995), each session started by explaining how the auction process works. Then two practice rounds were implemented to teach how the WTP auctions would be carried out. During those sessions, subjects submitted bids for a dollar bill and a chocolate bar

and became familiar with the bidding process used in the auctions for the chocolate.

During the chocolate auction, subjects were organized according to the previous randomized treatment design. For treatment group one (NO INFO), subjects were told to taste three different types of chocolate. Right after each tasting, the subjects answered two questions regarding the taste and smell of each variety and were told to submit a bid from \$0 to \$10, reflecting their maximum willingness to pay for 3.5 ounces of the chocolate just tasted. At the end of the experiment, we randomly chose just one of the auctions to be binding, so at most, the subject only purchased 3.5 ounces of chocolate. Also, we randomly drew a market price for the 3.5 ounces of chocolate in the selected auction. Only those who bid equal to or above the market price won the auction. The winners received the chocolate plus the difference (e.g., \$10 minus the market price); the rest received \$10. Finally, the subjects completed an exit survey regarding their demographics and chocolate consumption trends.

On the other hand, three additional sessions were conducted for treatment group two (INFO). The only difference between the groups was that for the latter, the subjects were told the nutritional attributes of each chocolate variety, such as the quality of the cocoa beans used during the manufacturing process and the percentage of daily potassium and total

daily fat per 3.5 ounces serving before they made their bid. Table 3.1 shows the information provided to the subjects regarding the three varieties of chocolate.

Table 3.1 Experimental design: Information provided for the three varieties of Ecuadorian chocolate

National Ecuadorian Chocolate	Standard Ecuadorian Chocolate (CCN-51)	Banana Chocolate (CCN-51 Banana)
Made with the highest quality cocoa beans	Made from commercial standard cocoa beans	Made with standard cocoa beans and banana
0% of daily potassium per 3.5 ounce serving	0% of daily potassium per 3.5 ounce serving	26% of daily potassium per 3.5 ounce serving
54% of daily total fat per 3.5 ounce serving	54% of daily total fat per 3.5 ounce serving	19.5% of daily total fat per 3.5 ounce serving

Source: Authors' elaboration

We are interested in the differences in WTP between the National, CCN-51, and CCN-51 Banana varieties for the two treatment sources. For instance, the variety CCN-51 is the one made from the CCN-51 hybrid cocoa, so its comparison against the National variety will tell us if the consumers penalized the former for its lower quality compared to the latter, which in turn might help us to determine if the chocolate made from CCN-51 is a good substitute for the National variety (which is highly appreciated in the international markets). Also, comparisons among treatment groups will help us determine the impact of the information provided in the WTP for each variety.

3.4 Data and Empirical Specification

We collected 336 observations from 112 student subjects participating in the chocolate auctions. Each subject participated in auctions for 3.5 ounces of three chocolate varieties. 57 subjects were part of treatment group 1 (NO INFO), and 55 were part of treatment group 2 (INFO). In each auction, participants submitted bids, between \$0 and \$10 for 3.5 ounces of each chocolate variety. After all bids were submitted, participants completed a survey regarding demographic and purchasing habit information, including gender, age, education, income, chocolate consumption, and whether the participant is the primary shopper in their family or not.

Table 3.2 provides descriptive statistics illustrating the WTP for the different varieties of Ecuadorian Chocolate and the two treatment groups. The table also depicts demographic and purchase habit information about the subjects. The table shows that the average bid was \$2.26 per 3.5 ounces. For both treatments (NO INFO and INFO), the highest bid was for the National (highest quality), followed by the CCN-51 and the CCN-51 banana varieties (see Figure 1). However, in the treatment group INFO, which received information about the quality of the cocoa used as presented and some nutritional facts about the chocolate as presented in Table 3.1, the mean bid for the National and the CCN-51 varieties

decreased compared to the treatment group NO INFO (from \$2.75 to \$2.65 for the National and \$2.65 to \$2.47 for the CCN-51).

Table 3.2 Descriptive statistics for Willingness to Pay (WTP), demographics, and habit purchase Data

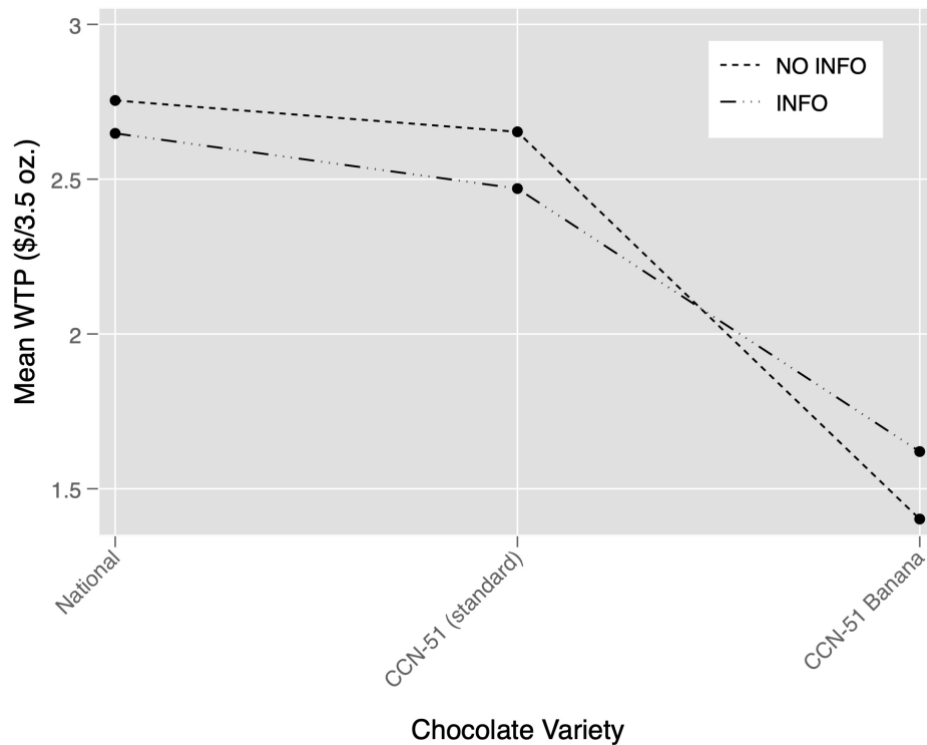
Variable	N	Mean	Standard Deviation	Minimum	Maximum
<i>WTP (\$ per 3.5 ounces)</i>					
<i>Treatment 1: NO INFO</i>	171	2.27	1.81	0.00	9.00
National	57	2.75	1.63	0.50	8.00
CCN-51	57	2.65	1.84	0.10	9.00
CCN-51 Banana	57	1.40	1.67	0.00	7.00
<i>Treatment 2: INFO</i>	165	2.25	1.85	0.00	10.00
National	55	2.65	2.05	0.00	10.00
CCN-51	55	2.47	1.81	0.50	10.00
CCN-51 Banana	55	1.62	1.51	0.00	5.00
<i>Overall</i>	336	2.26	1.83	0.00	10.00
<i>Demographic variables</i>					
Gender	112	0.51	0.50	0.00	1.00
Age	112	0.06	0.24	0.00	1.00
Education	112	0.16	0.36	0.00	1.00
Income	112	0.20	0.40	0.00	1.00
<i>Consumer trends variables</i>					
<i>How often check nutritional facts</i>					
Never	112	0.23	0.42	0.00	1.00
Sometimes	112	0.49	0.50	0.00	1.00
Always	112	0.28	0.45	0.00	1.00
<i>How often eat chocolate</i>					
Less than once a week	111	0.47	0.50	0.00	1.00
1-3 times a week	111	0.28	0.45	0.00	1.00
more than 3 times a week	111	0.25	0.44	0.00	1.00
<i>% of purchased chocolate is</i>					
highest quality	111	45.54	33.34	0.00	100.00
Primary Shopper in house	111	0.22	0.41	0.00	1.00

Notes: The table shows descriptive statistics for the variables used in the analysis. People in the treatment INFO were provided information about the different varieties of chocolate as described in Table 3.1. Gender is a dummy = 1 if female. Age is a dummy = 1 if the subject is 24 years or older. Education is a dummy = 1 if the subject has a college or more education. Income is a dummy = 1 if income >\$20,000. Source: Authors' elaboration.

In contrast, for the CCN-51 Banana, the mean bid increased (from \$1.40 to \$1.62). This result is consistent with the literature that consumers prefer chocolate low in fat and rich in nutrients (Visschers & Siegrist, 2009).

The bottom portion of Table 3.2 shows the responses to the demographic questions included in the survey. 51% of the subjects in our sample are female. 94% of the subjects are between 18 and 24 years old. 16% have college or more education. 80% of the subjects earn less than \$20,000 a year. This is expected since most of the subjects are undergraduate students. 72% of the subjects rarely or sometimes check nutritional facts, and 47% consume chocolate less than once a week. On average, 45% of the chocolate purchased by the subjects is of the highest quality. Finally, only 22% of the subjects are primary shoppers at their homes.

Figure 3.1 Comparison of WTP among treatment groups



Notes: The figure depicts the willingness to pay for 3.5 oz on the y-axis across different varieties of chocolate and treatment regimens according to Table 3.1. Source: Authors' elaboration.

Table 3.3 shows the balance in the explanatory variables across treatments. All the explanatory variables are balanced except the one related to income and the dummy for whether subjects always check nutritional facts where there are statistically significant differences in means at a 5% significance level. The difference in the income variable is explained by the fact that few subjects earn more than \$20,000 a year (22)⁸⁰, and most ended up in the INFO treatment (17). Something similar happened for the dummy on whether subjects always check nutritional facts. Only 20% of subjects always check for nutrition facts (31), and most ended up in the INFO treatment (20). We can control for these demographic and purchase habit differences when presenting our main estimation results (see equation 1).

Table 3.3 Balance in the explanatory variables. Mean differences by information treatment across explanatory variables

	Treatment		Difference	
	No Info	Info		
<i>Demographic variables</i>				
Gender	0.49	0.53	-0.04	
Age	0.05	0.07	-0.02	
Education	0.18	0.15	0.03	
Income	0.09	0.31	-0.22	*
<i>Consumer trends variables</i>				
<i>How often check nutritional facts</i>				
Never	0.26	0.20	0.06	
Sometimes	0.54	0.44	0.11	
Always	0.19	0.36	-0.17	*
<i>How often eat chocolate</i>				
Less than once a week	0.48	0.45	0.03	
1-3 times a week	0.29	0.27	0.01	
more than 3 times a week	0.23	0.27	-0.04	
% of purchased chocolate is				
highest quality	48.45	42.60	5.85	
Primary Shopper in house	0.25	0.18	0.07	

⁸⁰ This is expected since most of the subjects are undergraduate students.

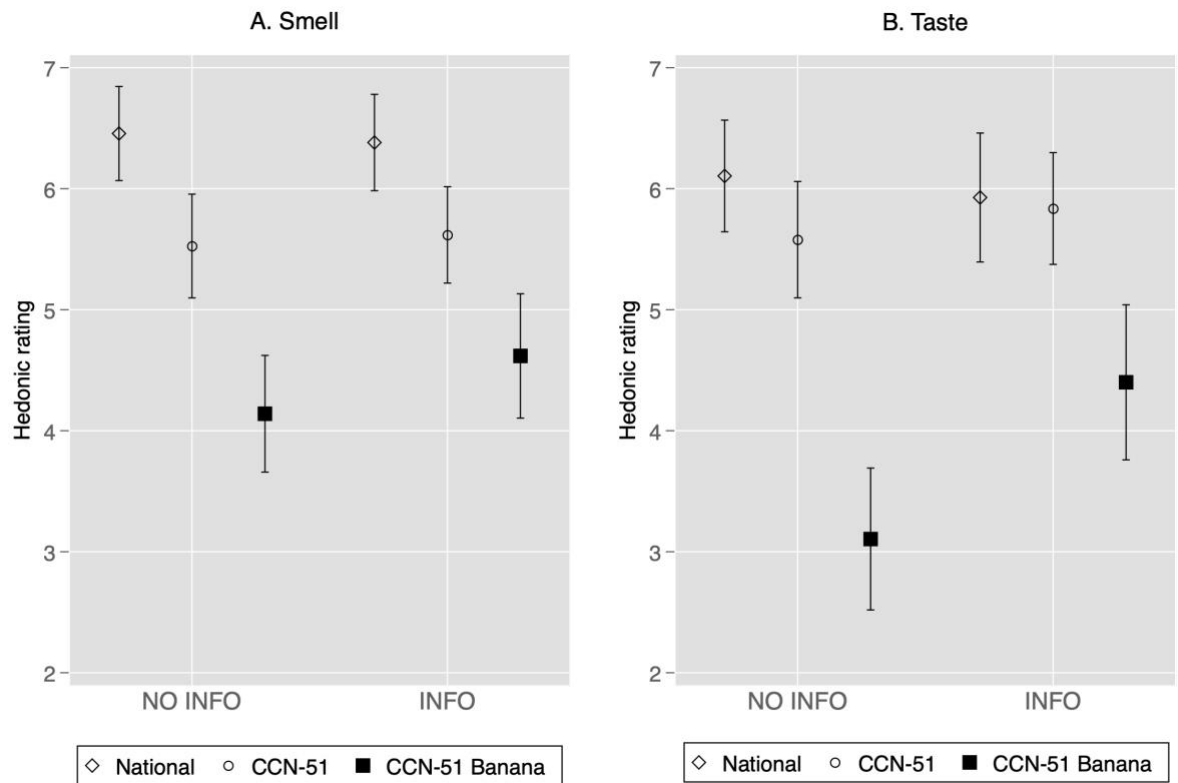
Notes: The table shows the mean differences between the explanatory variables across information treatments and a t-test for the difference in means. All the explanatory variables are balanced except the one related to income and the dummy for whether subjects always check nutritional facts where there are statistically significant differences in means at a 5% significance level. The difference in the income variable is explained by the fact that few subjects earn more than \$20,000 a year (22), and most end up in the INFO treatment (17). Something similar happened for the dummy on whether subjects always check nutritional facts. Only 20% of subjects always check for nutrition facts (31), and most ended up in the INFO treatment (20). * Significant at 5%. Source: Authors' elaboration.

During the first stage of the auction process, the subjects were also asked to rank the taste and the smell on a scale from 1 (worst) to 9 (best). Figure 2 shows confidence interval plots for the hedonic ratings, smell (panel A) and taste (panel B), given (y-axis) across different varieties of chocolate and information treatment treatments (x-axis). The results show that the variety with the highest average ratings for both smell and taste and across treatments was the National, followed by the CCN-51 and the CCN-51 Banana.

However, the differences were less marked for taste than for smell ratings when comparing the National vs. the CCN-51. Meaning that independent of information treatment, subjects rated the smell of the National variety as superior to the CCN-51 (see Figure 2, Panel A). This was not the case for the taste ratings once information was provided (see Figure 2, Panel B). For instance, taste ratings for the CCN-51 variety increased once information was provided while slightly decreasing for the national variety to the point that there were no meaningful mean differences in taste once information was provided. Meaning that once

information on the quality of the beans was revealed, subjects not only did not pay a premium for the supposedly highest quality (e.g., National), but they did pay slightly more for the chocolate made from standard cocoa beans (e.g., CCN-51).⁸¹

Figure 3.2 Hedonic properties of the different varieties of chocolate by treatment group



Notes: The figure depicts the mean hedonic ratings (smell and taste) and their 95% confidence intervals for the three different varieties of chocolate across the two treatment sources, as described in Table 3.1. The graphs were created using the package *ciplot* in Stata 17. Source: Authors' elaboration.

⁸¹ This result seems counterintuitive and could be explained by several students being confused about the sources of treatment in a way that they assigned relatively higher ratings to the variety that they thought were of the highest quality when in fact it was not. However, we don't observe particular individuals in the data that dramatically shift their ratings once the information is provided.

Finally, the trends in smell and taste ratings were similar for the CCN-51-Banana variety once information on banana, fat, and potassium content was disclosed: ratings tended to increase once information was revealed. However, the effect of information was higher for taste than the smell. For instance, the increase in smell and taste ratings once information was provided was 0.47 and 1.3 units, respectively. These results are in line with previous literature. When information about reduced fat content is provided, there is an increase in the acceptance of reduced-fat products (Kahkonen et al., 1996).

Table 3.4 shows the results of repeated measures ANOVA⁸² for the hedonic ratings (smell and taste) and the willingness to pay (\$ per 3.5 ounces) by information treatment. Since the ANOVA analysis only determines whether the difference among groups exists, but not which groups are different from each other, we complement the ANOVA results with post hoc Tukey Tests for pairwise comparisons. The ANOVA results show differences among the three chocolate varieties for the smell and taste ratings. However, the individual pairwise comparisons show that these differences follow different patterns for smell and taste ratings. On the one hand, for the smell ratings, there are differences between all

⁸² Repeated measures ANOVA account for the fact that we are interested in differences in means between groups that are not independent. For instance, in this document, we are interested in differences in mean hedonic ratings and WTP for the same individuals under different information conditions. Therefore, differences between hedonic ratings and WTP among different varieties depend not only on comparing the smell and taste profiles, but also on the information treatment given.

groups (National vs. CCN-51; National vs. CCN-51 Banana; and CCN-51 vs. CCN-51 Banana). All the comparisons suggest that the National chocolate variety is the highest rated, followed by the CCN-51, and CCN-51 Banana varieties. The mean difference between the national and the CCN-51 varieties is 0.76 units in the treatment group with information and 0.93 units among the subjects who did not receive information.

On the other hand, taste ratings results show no differences in the mean taste ratings between the CCN-51 and national chocolates. The result is similar in both the INFO and not INFO treatments. Indeed, mean differences in WTP decrease once information is provided from \$0.53 to \$0.09 per 3.5 oz. This result is highly relevant for the stakeholders of the chocolate industry because it demonstrates that consumers do not consider the taste of the chocolate made with CCN-51 far away from the national variety, even when provided information about the highest quality of the National variety. Furthermore, it also suggests that the new fermentation techniques applied to the CCN-51 might improve its hedonic properties and open the door for this hybrid to play a higher regarded role in the industry.

Table 3.4, Panel B, also shows the same analysis for the willingness to pay variable. The results suggest the existence of differences in WTP among different varieties of Chocolate for both information treatments.

However, the pairwise Tukey comparisons show that these differences are only present when we compare the National against the CCN-51 Banana variety and the CCN-51 against the CCN-51 Banana variety. There is no difference in WTP when we compare the CCN-51 vs. the National variety. This result holds for both treatment information groups. Again, these results might suggest that despite the CCN-51 variety chocolate being, theoretically, from less quality cocoa, the consumers do not perceive those differences in hedonic properties (at least for the taste) to be significant enough to elicit significant differences in WTP. In line with previous literature, information on health-related attributes of chocolate has little effect on chocolate consumer behavior. Consumers still pay more for chocolate with better taste, even in the presence of health-related information (Steinhauser et al., 2019).

Table 3.4 Differences in hedonic ratings and WTP among chocolate varieties by information treatment

<i>Panel A: ANOVA results for hedonic ratings</i>				
	Smell		Taste	
	INFO	NO INFO	INFO	NO INFO
Differences among groups	✓	✓	✓	✓
Tukey HSD tests	Mean differences			
CCN-51 vs. National	-0.76 [-1.39, -0.13]	-0.93 [-1.53, -0.32]	-0.091 [-0.82, 0.64]	-0.53 [-1.32, 0.27]
CCN-51 Banana vs. National	-1.76 [-2.39, -1.13]	-2.32 [-2.92, -1.71]	-1.53 [-2.26, -0.80]	-3.00 [-3.79, -2.20]
CCN-51 Banana vs. CCN-51	-1.00 [-1.63, -0.37]	-1.39 [-1.99, -0.78]	-1.44 [-2.16, -0.71]	-2.47 [-3.27, -1.68]
<i>Panel B: ANOVA results for willingness to pay</i>				
	WTP			
	INFO	NO INFO		
Differences among groups	✓	✓		
Tukey HSD tests	Mean differences			
CCN-51 vs. National	-0.18 [-0.68, 0.32]	-0.10 [-0.60, 0.40]		
CCN-51 Banana vs. National	-1.02 [-1.53, -0.52]	-1.35 [-1.85, -0.85]		
CCN-51 Banana vs. CCN-51	-0.85 [-1.35, -0.35]	-1.25 [-1.75, -0.75]		

Notes: The table shows the results of repeated measures ANOVA tests and a post hoc Tukey Test for the differences in hedonic properties (Panel A) and WTP (Panel B) between varieties of Chocolate for the two sources of information as described in Table 3.1. The check symbol (✓) on the respective rows indicates the existence of the differences on the omnibus statistical ANOVA test to detect differences in groups. The null hypothesis of the ANOVA test for each hedonic rating and WTP is $H_0: \mu_{National} = \mu_{CCN-51} = \mu_{CCN-51-Banana}$. The Tukey HSD tests after ANOVA identify where the differences between the pair of means (mean ratings or mean WTP) are significant while controlling for the multiple group testing and experiment error rate (probability of a type I error). 95% confidence intervals for the mean difference in brackets. Source: Authors' elaboration.

The ANOVA results are an essential step in the identification process. Nonetheless, some heterogeneous preferences could emerge, and these could be guided by the demographic or purchase characteristics of

the consumers. We use a Tobit specification to model the WTP. A linear model for WTP is not adequate because WTP data is censored from below at zero ($WTP \geq 0$) (Wooldridge, 2010). The latent value of WTP for variety j in information treatment t for individual i , denoted as WTP_{jti}^* , is expressed as a function of the chocolate variety V_j , the information treatment for the two new varieties I_t , and the subjects' demographic and purchasing habits characteristics X_i . Because individuals submitted bids for different chocolates in the experiment, we employ a random-effects Tobit model to account for the panel nature of the data. The parameter v_i is an individual-specific disturbance for subject i , and ε_{jti} is the error term assumed to follow a normal distribution with mean zero and standard deviation σ . In Equation (1), we assume a linear functional form for the WTP equation. However, we observe the variable WTP_{jti} that is related to WTP_{jti}^* , and the relationship between the two is shown in Equation (2).

$$WTP_{jti}^* = \alpha + \beta V_j + \gamma I_t + \delta V_j I_t + \theta X_i + v_i + \varepsilon_{jti} \quad (1)$$

$$WTP_{jti} = \max\{0, WTP_{jti}^*\} \quad (2)$$

In the model specified above, α is the bid for 3.5 ounces of the CCN-51 chocolate variety when no information about the origins of the three types of chocolate is revealed, β captures the price differences that consumers are willing to pay for the National and CCN-51 Banana varieties compared to the CCN-51 variety, γ is a vector of parameters

controlling for the information treatments, δ is a vector of parameters describing the interaction effects between chocolate varieties and information treatments, and these capture the price premium that consumers are willing to pay for the two new varieties under different information treatments, and θ is a vector of parameters capturing how consumer characteristics affect the WTP for the different varieties of chocolate.

3.5 Results and Discussion

Figure 3.3 presents the results for two specifications: The Pool OLS and Random Effects Tobit Estimators. The results from the OLS model are close to those from the Tobit model. These results are difficult to interpret, given the number of interactions and categorical variables used. Initially, for the sake of simplicity, we discuss the variables that significantly affect the WTP. On the one hand, using the results from the Tobit model, we can determine that for one unit increase in the taste ratings, the consumers are willing to pay, on average, \$0.44/3.5 oz. On the other hand, the smell does not seem to play a role in the willingness to pay. This result is in line with previous literature where taste is the main contributing factor when choosing what chocolate to consume (Ozretic-Dosen et al., 2007; Ahmed et al., 2012), and even in situations where ethical claims related to fairtrade are also offered (Poelmans & Rousseau, 2016).

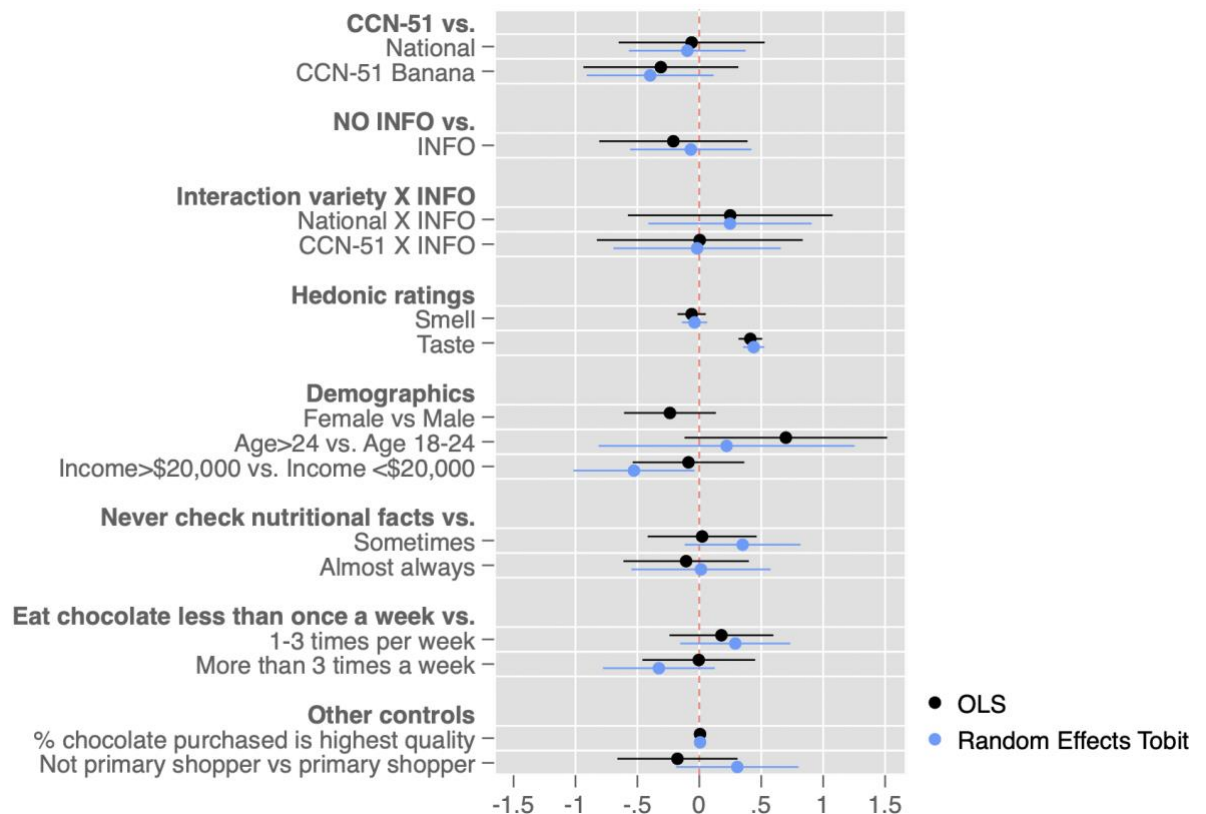
There was no significant difference in willingness to pay by gender, income, and age. Other variables related to customers' purchasing behavior, such as whether customers check nutritional facts, the frequency of chocolate purchasing, and the preference for high-quality chocolate, also seem to play a minor role in explaining the WTP in our sample. This result contrasts with previous literature in which women's WTP for Chocolate is higher and income positively affects WTP (Vecchio & Annunziata, 2015). Our results for age and income need to be interpreted with caution since most of the participants in the experiment are college students aged 18 to 24 years old (95%) with low income (91% of respondents earn less than \$20,000 a year). Therefore, any statistical test will not have enough power to detect differences by income and age.

In this document, we are particularly interested in determining if there are differences in WTP between the national chocolate and the CCN-51 variety. We also want to determine if these differences are conditioned by the treatment (INFO vs. NO INFO). To simplify those comparisons, we use the Tobit model coefficients (from Figure 3) to compute the average adjusted predictions (AAPs) for the WTP for each chocolate variety and treatment combination. Then we test for the existence of differences between AAPs⁸³. Table 3.5 reports the results of this exercise. Panel A in

⁸³ For instance, the AAPs for the treatment (INFO) variable are calculate as follows: first, go to the first case and treat that person as though she/he were treated (e.g., as he/she

Table 3.5 shows that the national variety has the highest predicted mean WTP at \$2.33, followed by the CCN-51 at \$2.30 and the CCN-51 Banana at \$1.90. There is no significant difference in mean WTP by treatment level: \$2.17 with NO INFO vs. 2.18 with INFO.

Figure 3.3 Willingness to Pay Estimates Using Random-effects Tobit and OLS Models



Notes: Figure depicts the estimation results for Equation 1 for two specifications: OLS and Random Effects Tobit (RE Tobit). The Chi-square statistics for an LR test statistic comparing the pooled Tobit estimator vs. the panel Tobit estimator is 49.81. The results suggest that the contribution of the panel level is big enough to use the random effects Tobit model. Standard errors in the OLS model cluster at the individual level.

receive information), regardless of what the treatment was. Hold all other independent variable values constant. Second, compute the prediction of the WTP for this person. Third, repeat 1 and 2 for all the subjects in the sample. Finally, compute the average of all predictions.

On the other hand, the interaction between the variety and treatment variables provided interesting results. Panel B in Table 3.5 presents the marginal effect of receiving INFO vs. NO INFO on the Adjusted prediction for WTP for each chocolate variety. When information is provided, the WTP for the national variety increases by \$0.18. Meanwhile, the WTP for the CCN-51 and the CCN-51 Banana decreases by \$-0.07 and \$0.09, respectively. However, the marginal effect of having information does not significantly impact the mean differences in predicted WTP. In a nutshell, this means that information about the chocolate varieties' health does not significantly affect the WTP of the subjects.

Finally, in the bottom part of Table 3.5, we compare the mean adjusted WTP predictions between the national and CCN-51 varieties. We believe this result might convey important information about how the subjects value the highest quality chocolate (national) compared with the one made from CCN-51. Results show no difference in mean WTP between these varieties holding the rest of the variables constant. This holds for both treatment groups. Taken at face value, this result might imply that there is an opportunity to substitute national cocoa with CCN-51 cocoa in the production of chocolate without hurting the producers. Therefore, it places CCN-51 cacao as an element to consider in dealing with the potential shortage of chocolate.

Table 3.5 Average Adjusted Predictions (AAPs) by Chocolate Type and Treatment

Variables	Average Adjusted Predictions (AAPs)	SE	95% CI
<i>Panel A: AAPs by Chocolate Variety and Treatment</i>			
National	2.33	0.19	[1.96, 2.69]
CCN-51	2.30	0.18	[1.94, 2.66]
CCN-51 Banana	1.90	0.19	[1.51, 2.27]
<i>Treatment</i>			
NO INFO	2.17	0.17	[1.83, 2.51]
INFO	2.18	0.17	[1.84, 2.52]
<i>Chocolate Type × Treatment</i>			
National × NO INFO	2.24	0.23	[1.79, 2.68]
National × INFO	2.42	0.22	[1.97, 2.86]
CCN-51 × NO INFO	2.33	0.22	[1.90, 2.77]
CCN-51 × INFO	2.27	0.22	[1.83, 2.70]
CCN-51 Banana × NO INFO	1.94	0.25	[1.47, 2.41]
CCN-51 Banana × INFO	1.85	0.23	[1.40, 2.30]
<i>Panel B: Effect of Treatment in WTP</i>			
	<i>Marginal effect</i>	SE	95% CI
(National × NO INFO) - (National × INFO)	0.18	0.25	[-0.42, 0.78]
(CCN-51 × NO INFO) - (CCN-51 × INFO)	-0.07	0.25	[-0.66, 0.53]
(CCN-51 Banana × NO INFO) - (CCN-51 Banana × INFO)	-0.09	0.26	[-0.71, 0.54]
<i>Panel C: Pairwise Comparisons, National vs. CCN-51</i>			
	P>chi2		
(National × NO INFO) = (CCN-51 × NO INFO)	0.69		
(National × INFO) = (CCN-51 × INFO)	0.53		

Source: Authors' elaboration

3.6 Conclusion

This paper explored the potential for Chocolate made with CCN-51 cocoa to replace chocolate made with traditional cocoa. We used an auction experiment with Cornell University students to test differences in consumers' WTP for three chocolate varieties, two using CCN-51 cocoa. We included different information treatments in the experiment to assess how the information provided affects consumers' WTP for the chocolate varieties. We applied a univariate analysis (ANOVA and Tukey Pairwise comparison) and a Random Effects Tobit estimator to account for the censored nature of the WTP data.

Results show that taste is the main contributing factor in eliciting differences in WTP between chocolate varieties. Results also show no difference in WTP when we compare the National (highest quality) vs. the Standard CCN-51 variety. This result holds for both treatment information groups. This result has important policy implications. It might suggest that despite the standard variety of chocolate being from a less quality cocoa (made with CCN-51 cocoa), the consumers do not perceive that differences in hedonic properties (at least for the taste) are big enough to make a significant difference in WTP to elicit which suggest that the expansion in the use of CCN-51 cacao in chocolate production is a plausible alternative to deal with the potential shortage of chocolate.

However, our results need to be interpreted with caution since our sample and experimental design has some limitations. First, our sample is composed of mostly college students with low income, so additional empirical evidence with a broader sample is essential to solidify the findings of this document. Second, our experimental design does not provide information to participants on the GMO nature of the CCN-51 cocoa and therefore does not account for the fact that consumers, especially those who prefer organic chocolate options, tend to avoid GMO chocolate.

REFERENCES

- Ahmed, Z., Zbid, I., Sikander, A. & Noujaim, R. (2012). Does country of brand origin (COBO) matter for the Lebanese consumers? *EuroMed Journal of Business*, 7(2), 108-128.
- Amores, F.M., Vasco, S.A., Eskes, A.B., Suarez, C., Quiroz, J.G., Loor, R.G., Jimenez, J.C., Zambrano, J., Bolaños, M.J., Reynel, V.H., Teran, M.M., & G.C. Quijano. (2011). On-farm and on-station selection of new cocoa varieties in Ecuador, p. 59–72. In: Eskes, A.B. (ed.). Collaborative and participatory approaches to cocoa variety improvement. Final report of the CFC/ICCO/Bioversity International project on cocoa productivity and quality improvement: A participatory approach (2004–2010). Common Fund for Commodities (CFC), Amsterdam, The Netherlands; Intl. Cacao Organization (ICCO), London, UK; Bioversity Intl., Rome, Italy.
- Anand, A., Mittelhammer, R.C. & McCluskey, J.J. (2007). Consumer response to information and second-generation genetically modified food in India. *Journal of Agricultural & Food Industrial Organization*, 5(1), 1-18.
- Balineau G., Bernath, S. & Pahuatini, V. (2016). Cocoa farmers' agricultural practices and livelihoods in Côte d'Ivoire. AFD Technical Reports, No. 24.
- Barrientos, S. (2013). Gendered global production networks: Analysis of cocoa-chocolate sourcing. *Regional Studies*, 48(5), 791-803.
- Barrientos, S. (2016). Beyond fair trade: Why are mainstream chocolate companies pursuing social and economic sustainability in cocoa sourcing? In M. Squicciarini & J. F. Swinnen (Eds.), *The economics of chocolate* (pp. 213-228). Oxford: Oxford University Press.
- Becker, G., DeGroot, M. & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3), 226-232.
- Boza, E. J., Motamayor, J., Amores, F., Cedeño-Amador, S., Tondo, C., Livingstone, D., Schnell, R., & Gutiérrez, O. (2014). Genetic characterization of the cacao cultivar CCN-51: its impact and significance on global cacao improvement and production. *Journal of the American Society for Horticultural Science*, 139(2), 219-229.

- Curtis, K. R., McCluskey, J.J. & Wahl, T. I. (2004). Consumer acceptance of genetically modified food products in the developing world. *The Journal of Agrobiotechnology Management and Economics*, 7, 70-75.
- Dannenber, A., Scatasta, S. & Sturm, B. (2008). Does mandatory labeling of genetically modified food grant consumers the right to know? Evidence from an economic experiment. Centre for European Economic Research, discussion paper No. 08-029.
- IDH (2014), "Cocoa: strong business case creates unprecedented drive for mainstream market transformation," available at: www.idhsustainabletrade.com/cocoa (accessed July 21, 2017).
- ICCO (2017). Quarterly Bulletin of Cocoa Statistics, 43(2), Cocoa year 2016/17.
- ICCO (2022). Quarterly Bulletin of Cocoa Statistics, 48(3), Cocoa year 2021/2022.
- FAOSTAT (Food and Agriculture Organization of the United Nations Statistics Division). (2013). Production Statistics: Crops. Available at: <http://faostat3.fao.org/browse/Q/QC/E>
- Fold, N. & Neilson, J. (2016). Sustaining supplies in smallholder-dominated value chains: Corporate governance of the global cocoa sector. In M. Squicciarini & J. F. Swinnen (Eds.), *The economics of chocolate* (pp. 195-212). Oxford: Oxford University Press.
- Grand View Research (2021). *Chocolate Market Size, Share & Trends Analysis Report By Product (Traditional, Artificial), By Distribution Channel (Supermarket & Hypermarket, Convenience Store, Online), By Region, And Segment Forecasts, 2022 – 2030*. Available at: <https://www.grandviewresearch.com/industry-analysis/chocolate-market#>
- Gilbert, C. (2016). The dynamics of the world cocoa price. In M. Squicciarini & J. F. Swinnen (Eds.), *The economics of chocolate* (pp. 307-338). Oxford: Oxford University Press.
- Greenpeace International (2008). Say no to Genetic Engineering. Retrieved from: <http://www.greenpeace.org/international/campaigns/genetic-engineering>, 22/04/08


- Janssen, E. & Riera, O. (2016). Too hot to handle: The explosive growth of chocolate in India. In M. Squicciarini & J. F. Swinnen (Eds.), *The economics of chocolate* (pp. 419-438). Oxford: Oxford University Press.
- Kagel, J. (1995). Auctions: a survey of experimental research. In J. Kagel & A. Roth (Eds.), *The Handbook of Experimental Economics* (pp. 501-585). Princeton, NJ: Princeton University Press.
- Kähkönen, P., Tuorila, H. & Hannu, R. (1996). How information enhances acceptability of a low-fat spread. *Food Quality and Preference*, 7(2), 87-94.
- Kähkönen, P., Hakanpää, P. & Tuorila, H. (1999). The effect of information related to fat content and taste on consumer responses to a reduced-fat frankfurter and reduced-fat chocolate bar. *Journal of Sensory Studies*, 14, 35-46.
- Kajale, D. B. & Becker, T.C. (2014). Effects of information on young consumers' willingness to pay for genetically modified food: experimental auction analysis. *Ecology of Food and Nutrition*, 53, 292-311.
- Kass, K.P. & Ruprecht, H. (2006). Are the Vickrey auction and the BDM mechanism really incentive compatible? – Empirical results and optimal bidding strategies in cases of uncertain willingness to pay. *Bidding Strategies*, 58, 37-55.
- Kox, H. (2000). The market for cocoa powder, background paper for modeling and forecasting the market for cocoa and chocolate, prepared for the ministry of foreign affairs, the Netherlands. Amsterdam: Economic and Social Institute.
- Moramarco, E. & Nemi, L. (2016). Nutritional and health effects of chocolate. In M. Squicciarini & J. F. Swinnen (Eds.), *The economics of chocolate* (pp. 134-156). Oxford: Oxford University Press.
- Norton, J.E., Fryer, P.J. & Parkinson, J.A. (2013). The effect of reduced-fat labeling on chocolate expectations. *Food Quality and Preference*, 28, 101-105.
- Noussair, C., Robin, S. & Ruffieux, B. (2002). Do consumers no care about biotech foods or do they just not read labels? *Economic Letters*, 75, 47-53.

- Noussair, C., Robin, S. & Ruffieux, B. (2004). Do consumers really refuse to buy genetically modified food? *The Economic Journal*, 114, 102-120.
- Lusk, J. and Fox, J. (2003). Value Elicitation in Retail and Laboratory Environments. *Economics Letters*, 79, 27–34.
- Läderach, P., Martinez-Valle, A., Schroth, G. & Castro, N. (2013). Predicting the future climatic suitability for cocoa farming of the world's leading producer countries, Ghana and Côte d'Ivoire. *Climate Change*, 119, 841-854.
- Schatzker, M. (2014, November 14). To Save Chocolate, Scientists Develop New Breeds of Cacao. *Bloomberg*, Retrieved from <https://www.bloomberg.com/news/articles/2014-11-14/to-save-chocolate-scientists-develop-new-breeds-of-cacao>.
- Schroth, G., Läderach, P., Martinez-Valle, A., Bunn, C. & Jassogne, L. (2016). Vulnerability to climate change of cocoa in West Africa: Patterns, Opportunities and limits to adaptation. *Science of the Total Environment*, 556, 231-241.
- Steinhauser, J., Janssen, M. & Hamm, U. (2019). Consumers' purchase decisions for products with nutrition and health claims: What role do product category and gaze duration on claims play? *Appetite*, 141.
- Ozretic-Dosen, D., Skare, V. & Krupka, Z. (2007). Assessments of country of origin and brand cues in evaluating a Croatian, western and eastern European food product. *Journal of Business Research*, 60, 130-136.
- Ploetz R. (2016), The impact of diseases on cacao production: A global review. In Bryan A. Bailey & Lyndel W. Meinhardt (Eds), *Cacao diseases: A history of old enemies and new encounters*. Springer.
- Poelmans, E. & Rousseau, S. (2016). How do chocolate lovers balance taste and ethical considerations? *British Food Journal*, 118(2), 343-361.
- Poelmans, E. & Swinnen, J. (2016). A brief economic history of chocolate. In M. Squicciarini & J. F. Swinnen (Eds.), *The economics of chocolate* (pp. 11-42). Oxford: Oxford University Press.
- Phillips-Mora, W. & J. Wilkinson. (2007). Frosty Pod of Cacao: A Disease with a Limited Geographic Range but Unlimited Potential for Damage. *Phytopathology*, 97, 1644-1647.


- Rommens, C. M. (2010). Barriers and paths to market for genetically engineered crops. *Plant Biotechnology Journal*, 8, 101-111.
- Rozin, P., Levine, E. & Stoess, C. (1991). Chocolate craving and liking. *Appetite*, 1991, 17, 199-212.
- Schmitz, H. & Shapiro, H. (2012, February). The Future of Chocolate. *Scientific American*, 306(2), 60-65.
- Tagbata, D. & Sirieix, L. (2008). Measuring consumer's willingness to pay for organic and Fair-Trade products. *International Journal of Consumer Studies*, 32, 479-490.
- Tamru, S. & Swinnen, J. (2016). Back to the roots: Growth in cocoa and chocolate in Africa. In M. Squicciarini & J. F. Swinnen (Eds.), *The economics of chocolate* (pp. 439-457). Oxford: Oxford University Press.
- van der Kooij, S. (2013). Market study of fine flavor cocoa in 11 selected countries. Royal Tropical Institute. Retrieved from:
<http://www.cocoaconnect.org/sites/default/files/publication/Market%20study%20of%20fine%20flavour%20cocoa%20-%20revised%20version.pdf>
- Visschers, V. & Siegrist, M. (2009). Applying the evaluability principle to nutrition table information. How reference information changes people's perceptions of food products. *Appetite*, 52(2), 505-512.
- Wooldridge, J. M. (2010). *Econometric analysis of cross-section and panel data*. The MIT Press.

APPENDIX

A. Slides explaining the experiment to the NO INFO group




Chocolate Tasting Experiment




Welcome to an experiment in consumer decision making!

- Read the **instructions** carefully and refrain from communicating with other participants
- Do not use any software programs on your computer other than the spreadsheet we will use for our auctions
- Refrain from using your cell phones during the experiment
- Your participation in this experiment is **voluntary** and you can withdraw from this experiment at any time
- Please read and sign the **Consent Form** if you agree to participate in the experiment



Instructions



- You will be paid \$10 for participating in this experiment
- You will participate in a series of auctions for 3.5 ounces of various types of chocolate
- It is very important that you follow the instructions and ask any questions you may have
- In order to participate in the [auction](#) you need to familiarize yourself with the auction process



How Will the Auction Work?



- Each participant will submit a bid for a product
- Next, we will randomly select a **market price** for it
- Those who bid **equal or higher** than the market price will get the product being auctioned at the randomly drawn price rather than their bid-level
- Those who bid **less than** the randomly drawn market price will not get the product

Hypothetical One Dollar Bill Auction



- Press continue to see the “One Dollar Bill Auction” Screen
- Type your bid (**between 0 and 2 dollars**) for the One Dollar and click the button to submit your bid
- We will then randomly draw the market price between \$0 and \$2 for the One Dollar bill
- Only those who bid equal or above the market price will “win” the One Dollar Bill
- It is important to understand that in order to maximize your earnings you should bid the value of the product.

Hypothetical One Dollar Bill Auction




In order to maximize your earnings you should bid your true willingness to pay for the product. Here it should be obvious that your willingness to pay would be \$1.00, consider:

- If your bid is \$0.80 AND the random **market price** is \$0.90, then you would forego the opportunity to buy a \$1 bill for \$0.90
- If your bid is above \$1, say \$1.10, AND the market price is between \$1 and \$1.10, say \$1.05, then you will pay \$1.05 for \$1 bill.
- It is always in **your best interest** to place a bid equal to your true willingness to pay for the product being auctioned!!


Hypothetical Skittles Auction




- Please submit your bid (between \$0 and \$1) for the candy bar shown by the administrator
- Type your bid on the cell highlighted in yellow and click "Submit"
- Please wait while we randomly select the market price between \$0 and \$2
- Did you win the candy bar?




Auctions for 3.5 ounces of Chocolate



- You will taste a number of types of chocolate produced in Ecuador.
- After each tasting, you will respond to two taste questions about the chocolate
- Next, you will submit a bid (from \$0 to \$10) reflecting your **maximum willingness to pay** for the chocolate just tasted
- We will use the same type of auction as we did for the one-dollar bill and candy bar



Auctions for 3.5 ounces of Chocolate




- At the end of the experiment, we will randomly choose just one of the auctions to be binding so at most you will only purchase 3.5 ounces of chocolate
- Also, we will randomly draw a market price for the 3.5 ounces of chocolate in the selected auction
- The **winners** receive the chocolate plus the difference (\$10 minus the market price); the rest receive \$10
- Finally, you will complete an **exit survey** and your receipt before receiving your payment and/or chocolate




We Are Starting Now with the Ecuadorian Chocolate Auctions





Chocolate A



- Clear your palate by eating an unsalted cracker and taking a sip of water.
- Take a look and smell the chocolate before tasting.
- Place the chocolate in your tongue and observe its taste and texture.



Chocolate B




- Again, clear your palate by eating an unsalted cracker and taking a sip of water.
- Take a look and smell the chocolate before tasting.
- Place the chocolate in your tongue and observe its taste and texture.




Chocolate C



- Again, clear your palate by eating an unsalted cracker and taking a sip of water.
- Take a look and smell the chocolate before tasting.
- Place the chocolate in your tongue and observe its taste and texture.



Final Instructions




- Please wait that we randomly select one of the chocolate auctions and the market price for the 3.5 ounces of chocolate
- Please complete the exit survey before receiving the payment
- Please complete your receipt before waiting in line to receive your cash and/or chocolate
- Thank you for your participation!

B. Slides explaining the experiment to the INFO group




Chocolate Tasting Experiment




Welcome to an experiment in consumer decision making!

- Read the **instructions** carefully and refrain from communicating with other participants
- Do not use any software programs on your computer other than the spreadsheet we will use for our auctions
- Refrain from using your cell phones during the experiment
- Your participation in this experiment is **voluntary** and you can withdraw from this experiment at any time
- Please read and sign the **Consent Form** if you agree to participate in the experiment



Instructions



- You will be paid \$10 for participating in this experiment
- You will participate in a series of auctions for 3.5 ounces of various types of chocolate
- It is very important that you follow the instructions and ask any questions you may have
- In order to participate in the [auction](#) you need to familiarize yourself with the auction process



How Will the Auction Work?



- Each participant will submit a bid for a product
- Next, we will randomly select a **market price** for it
- Those who bid **equal or higher** than the market price will get the product being auctioned at the randomly drawn price rather than their bid-level
- Those who bid **less than** the randomly drawn market price will not get the product

Hypothetical One Dollar Bill Auction



- Press continue to see the “One Dollar Bill Auction” Screen
- Type your bid (**between 0 and 2 dollars**) for the One Dollar and click the button to submit your bid
- We will then randomly draw the market price between \$0 and \$2 for the One Dollar bill
- Only those who bid equal or above the market price will “win” the One Dollar Bill
- It is important to understand that in order to maximize your earnings you should bid the value of the product.

Hypothetical One Dollar Bill Auction



In order to maximize your earnings you should bid your true willingness to pay for the product. Here it should be obvious that your willingness to pay would be \$1.00, consider:

- If your bid is \$0.80 AND the random **market price** is \$0.90, then you would forego the opportunity to buy a \$1 bill for \$0.90
- If your bid is above \$1, say \$1.10, AND the market price is between \$1 and \$1.10, say \$1.05, then you will pay \$1.05 for \$1 bill.
- It is always in **your best interest** to place a bid equal to your true willingness to pay for the product being auctioned!!

Hypothetical Skittles Auction

- Please submit your bid (between \$0 and \$1) for the candy bar shown by the administrator
- Type your bid on the cell highlighted in yellow and click "Submit"
- Please wait while we randomly select the market price between \$0 and \$2
- Did you win the candy bar?





Auctions for 3.5 ounces of Chocolate



- You will taste a number of types of chocolate produced in a Ecuador.
- After each tasting, you will respond to two taste questions about the chocolate
- Next, you will submit a bid (from \$0 to \$10) reflecting your **maximum willingness to pay** for the chocolate just tasted
- We will use the same type of auction as we did for the one dollar bill and candy bar



Auctions for 3.5 ounces of Chocolate



- At the end of the experiment, we will randomly choose just one of the auctions to be binding so at most you will only purchase 3.5 ounces of chocolate
- Also, we will randomly draw a market price for the 3.5 ounces of chocolate in the selected auction
- The **winners** receive the chocolate plus the difference (\$10 minus the market price); the rest receive \$10
- Finally, you will complete an **exit survey** and your receipt before receiving your payment and/or chocolate



We Are Starting Now with the Ecuadorian Chocolate Auctions





Please Listen to the definitions for each type of Ecuadorian Chocolate.



Chocolate A National Ecuadorian Chocolate





- Made with the highest quality cocoa beans
- 0% of daily potassium per 3.5 ounce serving
- 54% of daily total fat per 3.5 ounce serving



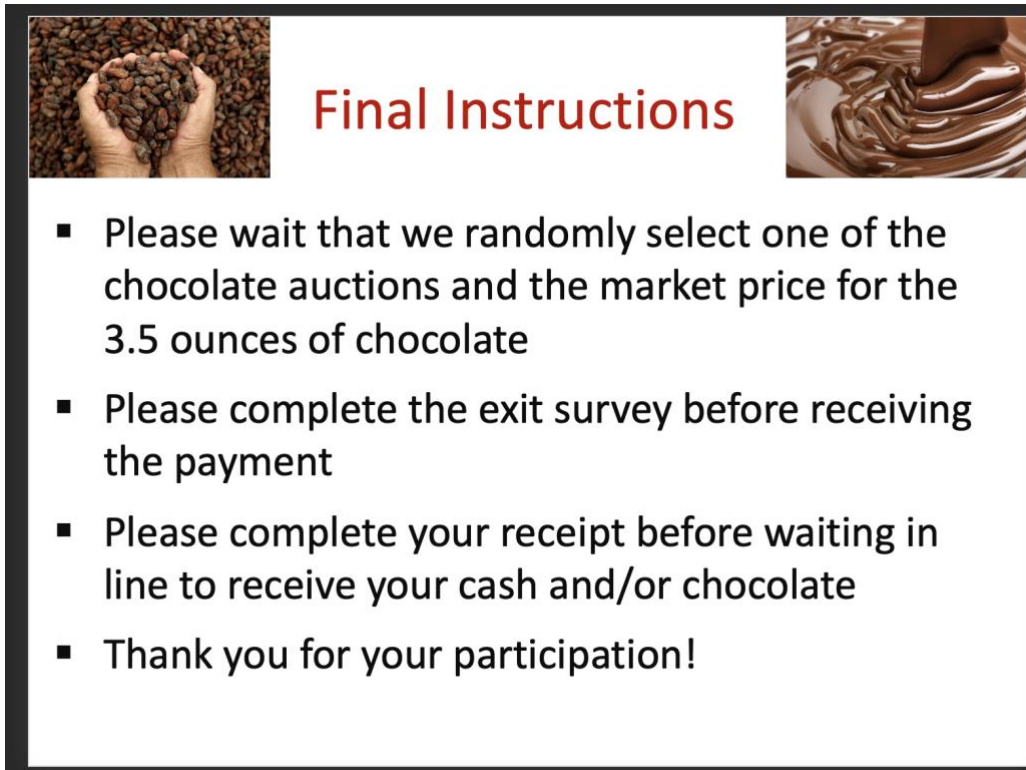
Chocolate B
Standard Ecuadorian
Chocolate

- Made from commercial standard cocoa beans
- 0% of daily potassium per 3.5 ounce serving
- 54% of daily total fat per 3.5 ounce serving



Chocolate C
Banana Chocolate

- Made with standard cocoa and banana
- 26% of daily potassium per 3.5 ounce serving
- 19.5% of daily total fat per 3.5 ounce serving



Final Instructions

- Please wait that we randomly select one of the chocolate auctions and the market price for the 3.5 ounces of chocolate
- Please complete the exit survey before receiving the payment
- Please complete your receipt before waiting in line to receive your cash and/or chocolate
- Thank you for your participation!

C. Consent Form

Estimating Willingness to Pay for Chocolate from different cocoa varieties. You are invited to take part in a research study that analyzes consumers' willingness to pay for chocolate from different types of cocoa. Please read this form carefully and ask any questions you may have before agreeing to take part in the study.

What the study is about: The purpose of this study is to understand some factors affecting consumers' willingness to pay for Chocolate from different cocoa varieties. Understanding these factors will help producers in developing countries make better production plans in order to increase their welfare.

What we will ask you to do: We will ask you to participate in a series of auctions for 3.5 ounces of different types of chocolate. You will be given basic information and a sample tasting of each type of chocolate and then be asked to submit your bid through a computer terminal. You will be endowed with \$10 to use for every bid. After all auctions have been completed, we will randomly select one auction that will be binding. That

means that, at most, you will purchase a bag of chocolate. Then, we will randomly select a number from 1 to 10 that will represent the market price. All bids equal to or above the randomly selected market price will purchase the chocolate. The persons who won in the selected auction will buy the chocolate at the selected market price and will keep the rest of their initial endowment. All others will keep their initial endowment but will not purchase the chocolate. At the end of the session, we will also ask you to complete a questionnaire which will elicit demographic and socioeconomic information regarding yourself.

Risks and benefits: We do not anticipate any risks to you participating in this study other than those encountered in day-to-day life.

The ingredients of the chocolates are: Cocoa solids, cocoa butter, chocolate liquor, sugar, and bananas. No nuts were included.

Compensation: There will be a monetary compensation of \$10.00 and extra points for marketing course.

Your answers will be confidential: The records of this study will be kept private. In any sort of report we make public we will not include any information that will make it possible to identify you. Research records will be kept in a locked file; only the researchers will have access to the records.

Taking part is voluntary: Taking part in this study is completely voluntary. You are free at all times to stop with the experiment.

If you have any questions: The researcher conducting this study is Professor Miguel Gomez. Please ask any questions you have now. If you have questions later, you may contact Rafael Ceden Paulson at rac425@cornell.edu. If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Institutional Review Board (IRB) at 607.255.5138 or access their website at <http://www.irb.cornell.edu>.

Statement of Consent: I have read the above information, and have received answers to any questions I asked. I consent to take part in the study.

Your Signature _____

Date _____

Your Name (printed)

Signature of person obtaining consent _____

Date _____

Printed name of person obtaining consent _____

Date _____

This consent form will be kept by the researcher for at least three years beyond the end of the study and was approved by the IRB on April, 2015.

D. Instructions Auction Experiment

Welcome to an experiment in consumer decision making. Please read these instructions carefully and refrain from communicating with other participants. Also, please refrain from using your cell phones during the experiment. As stated in the Consent Form, your participation in this experiment is voluntary and you can withdraw from this experiment at any time.

You are going to participate in auctions for 3.5 ounces of chocolate of different cocoa varieties (Chocolate A, B and C). It is very important that you follow the instructions and ask any questions you may have. In order to participate in the auction you need to familiarize yourself with the auction process first.

We will use an auction called a BDM auction. In this auction, each participant will submit a bid for a product within a specified range. Next we will randomly select a market price for the product. Those who bid equal or higher than the market price will win the auction and get the product being auctioned. They will pay the market price for the product (not their bidding price). The next two examples will help you familiarize with the BDM auction.

E. First Trial Auction: Bidding for a One Dollar bill

To better understand the auction process, you will participate in a hypothetical auction of One Dollar Bill.



Please indicate the maximum you would be willing to pay for the one dollar bill between \$0.00 and \$2.00 (with a maximum of 2 decimals) in the box below:

F. Second Trial Auction: Bidding for a Candy Bar

In this second hypothetical auction you will bid for Skittles bite size candies shown by the experiment administrator.



Please indicate the maximum you would be willing to pay for the Skittles between \$0.00 and \$2.00 in the box below:

G. Instruction-Chocolate Auctions

We will start to auction 3.5 ounces of chocolate from different cocoa varieties.

At the beginning of the auction we will display 3 types of chocolate to you, please have a close look at them and smell them. Next, you will taste the chocolate samples provide on your table and answer two questions about each. Then, you will be asked to place your maximum willingness to pay (between \$0.00 and \$10.00) for 3.5 ounces of each chocolate type.

The 3 types of chocolate are labeled as following:

- Chocolate A
- Chocolate B
- Chocolate C

After the auctions are completed, we'll randomly select the chocolate type that will be exchanged. Even though there will be three auctions conducted, only one will be selected for exchange. Hence, you will purchase at most 3.5 ounces of chocolate.

Next, we'll announce the market price for 3.5 ounces of chocolate for the type selected. The market price is a number randomly drawn from \$0.00 to \$10.00. Those who placed bids equal to or above the market price will win the auction and purchase the chocolate at the market price.

If you placed a winning bid in the selected auction (i.e. your bid was greater-or-equal than the market price), you will purchase 3.5 ounces of the product at market price and receive the balance in cash (\$10 minus the market price). If you did not win the selected auction, you will receive \$10 and the extra marketing points.

After the chocolate auctions are completed, you will be asked to complete a short survey. You will also be asked to complete a receipt to receive your payment and/or chocolate. Finally, you will proceed to the administrator to get your product and remaining balance.

Please rate the smell of each chocolate type: (from 1-9, with 9 being most favorable)

Chocolate A	<input type="radio"/> 9	<input type="radio"/> 8	<input type="radio"/> 7	<input type="radio"/> 6	<input type="radio"/> 5	<input type="radio"/> 4	<input type="radio"/> 3	<input type="radio"/> 2	<input type="radio"/> 1
Chocolate B	<input type="radio"/> 9	<input type="radio"/> 8	<input type="radio"/> 7	<input type="radio"/> 6	<input type="radio"/> 5	<input type="radio"/> 4	<input type="radio"/> 3	<input type="radio"/> 2	<input type="radio"/> 1
Chocolate C	<input type="radio"/> 9	<input type="radio"/> 8	<input type="radio"/> 7	<input type="radio"/> 6	<input type="radio"/> 5	<input type="radio"/> 4	<input type="radio"/> 3	<input type="radio"/> 2	<input type="radio"/> 1

Please rate the taste of each chocolate type: (from 1-9, with 9 being most favorable)

Chocolate A	<input type="radio"/> 9	<input type="radio"/> 8	<input type="radio"/> 7	<input type="radio"/> 6	<input type="radio"/> 5	<input type="radio"/> 4	<input type="radio"/> 3	<input type="radio"/> 2	<input type="radio"/> 1
Chocolate B	<input type="radio"/> 9	<input type="radio"/> 8	<input type="radio"/> 7	<input type="radio"/> 6	<input type="radio"/> 5	<input type="radio"/> 4	<input type="radio"/> 3	<input type="radio"/> 2	<input type="radio"/> 1
Chocolate C	<input type="radio"/> 9	<input type="radio"/> 8	<input type="radio"/> 7	<input type="radio"/> 6	<input type="radio"/> 5	<input type="radio"/> 4	<input type="radio"/> 3	<input type="radio"/> 2	<input type="radio"/> 1

Please indicate the maximum you would be willing to pay for 3.5 ounces of each chocolate type between \$0.00 and \$10.00 (with a maximum of two decimals) by entering the amount in the boxes below:

Chocolate A	Chocolate B	Chocolate C
\$	\$	\$

Please use the table below to help you calculate compensation:

Chocolate Type for Exchange:	<input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C
Market Price:	\$
Winning bid:	<input type="checkbox"/> Yes <input type="checkbox"/> No
Compensation: (please round decimals to the nearest whole number)	\$

H. Exit Survey

1. What is your gender?

Male Female

2. What is your age?

<input type="checkbox"/> Under 18 years	<input type="checkbox"/> 18-24 years	<input type="checkbox"/> 25-29 years
<input type="checkbox"/> 30-34 years	<input type="checkbox"/> 35-39 years	<input type="checkbox"/> 40-44 years
<input type="checkbox"/> 45-49 years	<input type="checkbox"/> 50-54 years	<input type="checkbox"/> 55-59 years
<input type="checkbox"/> 60-64 years	<input type="checkbox"/> 65-69 years	<input type="checkbox"/> 70-74 years
<input type="checkbox"/> 75-79 years	<input type="checkbox"/> 80 years or later	

3. What is the highest level of education you have completed?

<input type="checkbox"/> Less than High School	<input type="checkbox"/> High School/GED
<input type="checkbox"/> Some College	<input type="checkbox"/> 2-year College Degree
<input type="checkbox"/> 4-year College Degree	<input type="checkbox"/> Masters Degree
<input type="checkbox"/> Doctoral Degree	<input type="checkbox"/> Professional Degree (JD, MD)

4. What is your annual income range?

<input type="checkbox"/> Below \$20,000	<input type="checkbox"/> \$20,000-\$29,999	<input type="checkbox"/> \$30,000-\$39,999
<input type="checkbox"/> \$40,000-\$49,999	<input type="checkbox"/> \$50,000-\$59,999	<input type="checkbox"/> \$60,000-\$69,999
<input type="checkbox"/> \$60,000-\$69,999	<input type="checkbox"/> \$70,000-\$79,999	<input type="checkbox"/> \$80,000-\$89,999
<input type="checkbox"/> \$90,000 or more		

5. How often do you check nutritional facts when you buy a chocolate bar?

<input type="checkbox"/> Never	<input type="checkbox"/> Less than once a week
<input type="checkbox"/> 1-3 times a week	<input type="checkbox"/> 4-5 times a week
<input type="checkbox"/> More than 5 times a week	

6. How often do you eat chocolate?

Never

1-3 times a week

More than 5 times a week

Less than once a week

4-5 times a week

7. Please indicate approximately what percentage of chocolate that you purchase is regarded as highest quality chocolate (e.g. Lindt, Godiva, etc).

_____ %

8. Are you the primary shopper in your household?

Yes

No