ABSTRACT

Title of Dissertation:	ESSAYS ON POLITICAL ECONOMY OF DEVELOPMENT IN LATIN AMERICA
	Juan Carlos Angulo Santacruz Doctor of Philosophy, 2024
Dissertation Directed by:	Professor Erich Battistin Department of Agricultural and Resource Economics

This dissertation is composed of three applied economics essays in the intersection of development economics and political economy. The first provides an economic explanation to the increase in the intensity of conflict. The second focuses on the effects of presence of illegal activities on educational outcomes. The third analyzes the effect of mass migration on voting behavior and political preferences.

Chapter 1 studies how crime may be an unintended consequence of local development. The surge in global demand for Mexican avocados, particularly from the United States, has led to increased production and revenue in avocado-producing municipalities of Mexico. I show that these external changes in avocado global consumption patterns have influenced conflict levels in Mexico. Combining geographical variations in avocado cultivation suitability and fluctuations in avocado demand over time, I find a notable rise in homicides among agricultural workers in municipalities that are well-suited for avocado production. I demonstrate that this rise in homicides is not explained by the increased presence of drug cartels but rather heightened competition

between cartels for resources in municipalities where avocados are produced. These findings suggest that cartels vie for territorial control, diversifying their income sources, including the avocado industry, in response to their relatively limited influence over drug markets and routes.

In Chapter 2, I turn my attention to the production of illegal crops and how it affects schooling decisions. I focus on the case of Colombian coca leaves, the main input to produce cocaine. The country's main strategy to eradicate coca crops was the fumigation of herbicide until 2015, when the practice was banned. I exploit a plausible exogenous variation in the probability of being sprayed and the temporal effects of the fumigation campaigns as an instrument for the presence of coca fields. This temporal variation along with the cross-sectional variation of the spraying campaigns lead to an instrumental variable difference-in-differences. I use data on coca presence, eradication missions, and school outcomes at the municipal level from 2012 to 2018 to test whether a change in the presence of coca crops has an effect on schooling decisions. I show that my setting does not meet all the assumptions of the traditional difference-in-differences strategy but it fits those of Fuzzy Difference-in-Differences. My empirical findings suggest that an increase in the area cultivated with coca crops increases the high-school dropout rate and it has no effect on the enrollment rate. I rule out the possibility that coca presence crowds out other legal crops. Taken together, these results suggest that high school-age individuals are leaving school to work on coca related activities.

In Chapter 3, I revisit the question on whether political preferences of voters are molded by the presence of migrants. I exploit the unanticipated inflow to Colombia of Venezuelans fleeing their home country's political crisis in 2016 and the onset of economic collapse. I compare the results of the 2018 presidential campaign in Colombia across municipalities with similar trends in electoral outcomes between 2002 and 2014 but different presence of Venezuelan migrants on the verge of the 2018 campaign. To address the spatial sorting of migrants across these municipalities, I construct an instrumental variable based on the distance from the closest ports of entry. I find that an increase in the presence of migrants in the municipality yielded a polarized voting behavior. I show that these effects are explained by an increase in the electoral turnout, and that the fondness of voters for Colombia's 2016 Peace Agreement Plebiscite was an important determinant of their behavior, which has been overlooked in past empirical work.

ESSAYS ON POLITICAL ECONOMY OF DEVELOPMENT IN LATIN AMERICA

by

Juan Carlos Angulo Santacruz

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Erich Battistin, Chair/Advisor
Associate Professor Ethan D. Kaplan
Assistant Professor Cory B. Smith
Professor Anna Alberini
Assistant Professor Ramakrishna Sharan Mamidipudi

Acknowledgments

I am thankful to Erich Battistin, Ethan Kaplan, Cory Smith, Anna Alberini, and M.R. Sharan, my advisor and committee members, with whom I have had the chance to talk and learn from during various stages of my PhD. In addition, I would like to thank Allan Drazen and Isabella Alcañiz who played an important role on my training. I also owe thanks to other faculty outside of UMD who helped me to start this path, including Juan Fernando Vargas, Santiago Tobón, Chris Blattman, Leopoldo Fergusson, Andrés Carvajal, and Çağatay Kayı.

I would also like to thank my friends and colleagues from the Department of Economics and the Department of Government and Politics who supported me and encourage me to finish. A special mention to my Colombian friends who helped me through the hard times, and the Chilean community with whom I spent the pandemic. I can not list them all by name here as the list will be too long, however I would like to mention Camilo, Dario, Mayaris, Pilar, Catalina, and Pato.

I have presented these chapters at diverse seminars, workshops, and conferences. I am grateful to the discussants of my papers and with the people that took the time to listen, talk, or reach out to me with comments and suggestions. I have tried to the best of my abilities to include all the comments made to me during these events.

I acknowledge financial support from the Institute for Humane Studies and the government of Colombia through its program *Colfuturo*. While they provided support, I am the sole responsible for the content of this dissertation. All errors remain my own.

Table of Contents

List of Tables

List of Figures

Chapter 1: Green Gold: Avocado Production and Conflict in Mexico

1.1 Introduction

The expansion of high-value export crops in response to external demand shocks presents an important opportunity for poverty reduction and economic development (Boone, Kaila and Sahn, 2022; Uribe-Castro, 2021). However, for countries with low-quality institutions, this blessing may instead prove a curse, as an increase in the value of lootable resources may foster the emergence of criminal activities looking to profit from the booming sector (Cisneros, Hellmundt and Kis-Katos, 2022; Crost and Felter, 2019; Dube and Vargas, 2013; Mehlum, Moene and Torvik, 2006). These activities range from extortion to property and violent crimes to conflict incidents in civil wars (Charlton, James and Smith, 2021; Guardado and Pennings, 2020; Harari and Ferrara, 2018; Hastings and Ubilava, 2023; Ubilava, Hastings and Atalay, 2023) and lead to the emergence of rent-seeking criminal organizations, or the expansion of existing ones (Estancona and Tiscornia, 2022). For instance, the boom in the production of high-value exports at the end of the 19th century in Italy –sulfur, oranges and lemons– created a demand for private protection and led to the emergence of the Sicilian mafia (Buonanno et al., 2015; Dimico, Isopi and Olsson, 2017).

Given the increasing demand for high-value commodities and the persistent changes in

[0]This paper received funding from the Institute for Humane Studies, grant no. IHS017858.

economic conditions, how producers and various criminal organizations interact with each other remains an empirical puzzle. The increasing demand for high-value commodities could attract crime in simple forms, e.g., property or crop theft by individuals (Boone, Kaila and Sahn, 2022). However, it could also attract organized crime groups or criminal organizations, e.g., gangs, mafias, or drug trafficking organizations, that appropriate rents through extortion or violence (Dimico, Isopi and Olsson, 2017; Rios, 2013; Shirk and Wallman, 2015). Farmers may cooperate or coexist with criminal groups. In this case, criminal groups and farmers set an agreed upon 'tax' rate that farmers are willing to pay as quid pro quo for security and protection (Blattman, 2022; Buonanno et al., 2015). In this context, the presence of multiple competing criminal organizations could lead to confrontation over territorial control, with the winning group establishing a monopoly on looting (Atuesta and Pérez-Dávila, 2017; Blattman, 2022; Olson, 1993).

In this paper, I estimate the effects of increasing external demand for agricultural products on the intensity of conflict and whether the presence of criminal organizations further intensifies the conflict. The focus of my analysis is avocado production at the municipal level in Mexico during the period 2005-2019. As the largest producer and exporter of avocados in the world (Huang et al., 2023), as well as a home to an extensive network of criminal organizations (Atuesta and Pérez-Dávila, 2017), Mexico provides a compelling context in which to study the effects of demand shocks in agricultural production areas on conflict intensity. Avocados are Mexico's third most exported commodity and the United States its largest importer. The value of avocado exports increased over the time period covered by this study by over 455%. The increase in demand for avocados created a persistent and positive income shock that could affect conflict intensity and attract criminal organizations (Ferraz, Soares and Vargas, 2022). To the best of my knowledge, this paper is the first to empirically estimate the impact of a persistent income shock upon criminal behavior in an area with multiple criminal organizations.

The ideal experiment would compare otherwise identical municipalities that experience an increase in the production of avocados with those where production remains constant. In the absence of randomization, my analysis faces different challenges. First, the decision to plant and grow avocados is endogenous to farmers. Second, diverse agroclimatic conditions across different regions of the country introduce complexity, making it difficult to directly compare any two areas in terms of avocado production. Third, the decision to engage in avocado cultivation may exhibit correlations with socioeconomic and political conditions, increasing the likelihood of a municipality experiencing crime.

I tackle these challenges by considering a Bartik-type variable that combines geographical variation in avocado cultivation suitability with fluctuations in the United States' demand for avocados over time to create a measure of *potential avocado production* in a given year. This measure dynamically captures the areas more suited to meet the increased demand for avocados. Critically, I rely upon the identifying assumption that the United States' demand for avocados is not correlated with changes over time in violence across municipalities in Mexico (Goldsmith-Pinkham, Sorkin and Swift, 2020).

My empirical findings suggest that the rising demand for avocados increases the homicide rate at the municipal level. Using detailed information on the characteristics of the victims coming from death certificates, I show that this increase in the overall homicide rate is driven by an increase in the homicide rate of agricultural workers. Specifically, my estimates suggest that a one-standard-deviation increase in my measure of suitability to absorb the demand for avocado increases the homicide rate of agricultural workers by one additional death. This is a statistically and economically significant size effect of 20% when compared to the 5.78 mean homicide rate of agricultural workers. Furthermore, this increase explains 6% of the rise in violence, when compared to the 16.96 mean overall homicide rate in the sample.

My results are robust to using alternative measures of my main treatment variable, my outcome variables, and is also robust to estimation over different subsamples of my population. To validate my identification strategy, I also show that there is no relation between the homicide rate of agricultural workers and the suitability to grow crops other than avocados. Finally, to tackle some of the limitations of the fixed effects estimator with large longitudinal data (Millimet and Bellemare, 2023), I re-estimate my main specification using the Rolling First Difference (RFD) and the Rolling Twice-First Difference estimators, which accounts for the presence of unobserved time-varying heterogeneity. In particular, I show that my main estimates represent a lower bound for the actual treatment effect.

I use information on homicides, threats, incidents of extortion, and incidents of kidnapping from police reports to analyze effect upon other types of criminal activities. In terms of homicides, I show that my estimates are robust to the source of information about homicides. I find no changes in the number of threats or of kidnapping cases but that there is a decrease in the number of extortion cases in response to the demand shock. Although surprising, this last result can be explained by underreporting of criminal behavior to the police in the presence of increased cartel activity (Atuesta, Siordia and Lajous, 2018).

To examine the role of cartels, I use information from the Mapping Criminal Organizations project (Sobrino, 2020). First, I show that a one-standard-deviation increase in my measure of suitability to absorb the demand for avocados increases both the probability of cartel presence and the number of cartels in a given municipality. This increase is driven by both the expansion of traditional cartels—the 9 main Mexican drug cartels that have been present since the 1990s—and of the non-traditional cartels—the newly formed cartels or splinters of traditional cartels that have more local presence. Second, I show that the pre-existing presence of cartels, traditional or non-traditional, does not drive the increase in the homicide rate.

I also study cooperation between cartels and local communities, and competition between cartels. In terms of cooperation, I find a decrease in the homicide rate of agricultural workers in areas where cartels have operated since the 1990s, fully offsetting the increase given by the demand shock. This result lends credence to the idea that cartels may exert some form of governance over local communities (Blattman, 2022; Lessing, 2015, 2020; Sobrino, 2020). To proxy for territorial control, I study heterogeneous effects by presence of two or more cartels and find that the presence of multiple cartels –traditional or non-traditional– has no effect on homicides of agricultural workers. However, the presence of multiple newer, non-traditional cartels increases the homicide rate of cartel-related individuals. This result is consistent with the idea that criminal organizations battle to take control of a "profitable" territory in response to their relatively limited influence over drug markets and routes (Atuesta and Pérez-Dávila, 2017; Rios, 2013; Shirk and Wallman, 2015).

This paper makes three contributions to the resource curse and rent-seeking literature. First, to the best of my knowledge, this is the first paper to analyze a persistent income shock on conflict. Previous studies have focused on temporal income shocks that can be attributed to sudden and sharp short-lived changes in the price of the commodities (Boone, Kaila and Sahn, 2022; Cisneros, Hellmundt and Kis-Katos, 2022; Crost and Felter, 2019; Dube and Vargas, 2013; Dube, García-Ponce and Thom, 2016) or to changes in employment during harvesting season in labor intensive crops (Guardado and Pennings, 2020; Harari and Ferrara, 2018; Hastings and Ubilava, 2023; Ubilava, Hastings and Atalay, 2023). In these cases, demand for crops is more stable and the demand shock is less likely to transform conflict dynamics over an extended period.

Second, I show that increased demand for avocados increased the homicide rates of agricultural workers in municipalities more suitable for growing avocados. Using detailed information on the victims' characteristics, I can classify the homicide victims by occupation and show the presence of selective killings. Previous studies have not done this as they have used aggregated information on the overall homicide rate, violent incidents in civil conflict, such as attacks or clashes between groups (Crost and Felter, 2019; Dube and Vargas, 2013; Dube, García-Ponce and Thom, 2016; Guardado and Pennings, 2020; Harari and Ferrara, 2018; Ubilava, Hastings and Atalay, 2023), or data on social unrest, such as political violence, riots, and protests (Cisneros, Hellmundt and Kis-Katos, 2022; Crost and Felter, 2019; Hastings and Ubilava, 2023). The only exception I am aware of is the study of (De Haro, 2022), who uses a similar classification by occupation of homicides.

Third, I show how cooperation and competition between criminal organizations and farmers yield different results in terms of conflict intensity and victimization. In particular, I provide evidence that the presence of multiple criminal organizations leads to an increase in the homicide of cartel-related individuals, rather than the rest of the population. Other studies on competition have analyzed changes in the overall homicide rate when cartels enter or exit a municipality in Mexico (Sobrino, 2020), when the relative profitability of a crop increases, leading to competition between producers (De Haro, 2022; Dube, García-Ponce and Thom, 2016), or when new criminal opportunities emerge (Battiston et al., 2022; Estancona and Tiscornia, 2022; Piccato, 2022; Wainwright, 2016). These studies, however, have used aggregated measures of conflict that do not allow them to differentiate types of victims and may overestimate the effect of competition on homicides. The papers that are closer to mine are those of Estancona and Tiscornia (2022) and De Haro (2022). Estancona and Tiscornia (2022) develop a model suggesting that access to a good's export values creates opportunities for immediate profit, similar to the effect of the demand shock for avocados that I explore here, further affecting the intensity of the conflict. On the other hand, De Haro (2022) suggests that the increase in homicides in Mexico is driven by a substitution effect between avocado and poppy production. This latter set of results are similar to mine as she finds that areas more suited for growing avocados experience an increase in the homicide rate; however, she did not find that the demand shock attracts more presence of cartels in these areas. Her results complement mine as she focuses on the substitution effect whereas I focus on the persistent income effect of avocado production.

The rest of the paper is organized as follows. Section 1.2 presents the conceptual framework. Section 1.3 provides some institutional background. In section 1.4, I describe the empirical strategy to estimate the relation between the areas most suited for growing avocados and conflict, and section 1.5 explains the main sources of data. I present the main results of avocado production on homicides in section 1.6; while in section 1.7, I present robustness checks. In section 1.8, I explore the potential mechanisms that may explain the changes in homicides. Finally, the last section concludes.

1.2 Conceptual framework

To understand the relationship between economic shocks (e.g., agricultural production) and the intensity of conflict (e.g., homicides), I start by differentiating two main channels, as discussed in the literature. A positive economic shock should reduce crime by reducing the opportunity cost of engaging in criminal activities. However, certain types of shocks incentivize predation (Dube and Vargas, 2013; Ferraz, Soares and Vargas, 2022). Note that this analysis does not rely upon the presence of criminal organizations though criminal organization presence may accompany and accentuate predation.

Assume that criminal groups or cartels are profit-maximizing and follow a multi-commodity criminal approach, i.e., they profit from both legal and illicit industries through theft and extortion (Estancona and Tiscornia, 2022; Piccato, 2022; Wainwright, 2016). In the simplest case, assume two participants where each has one unit of resource. They can allocate any share of the resource into production, conflict –by attacking the other participant or defending themselves from an attack–, or no action (Dixit, 2004; Hirshleifer, 1991, 2001).[1] To simplify the analysis, assume participant A is a farmer and participant B is a cartel. Participant A is interested in production, but she can devote (part of) her time to defending herself from the attacks of participant B. Therefore, participant A will produce as long as either she can defend herself from the attacks of participant B or she comes to an agreement with participant B about how to share the output (Olson, 1993).

Assume that there is a technology of fighting such that there are returns to fighting (defense and offense), and there is a measure of efficacy of the offense over defense (Dixit, 2004; Hirshleifer, 1991, 2001). If participant A technology of fighting favors her defense over the offense of participant B, then the best action of participant B is not to attack; otherwise, participant B decision is to attack participant A to get part of her production. Suppose participant A increases her production and her technology of fighting favor does not favor the defense, then participant B will decide to attack to get a higher payoff. This leads to my first hypothesis:

[1] Notice, this set of actions is plausible in the context of Mexico. There is evidence of vigilantism, i.e., agricultural farmers creating self-defense groups to defend themselves from cartel attacks (Althaus and Dudley, 2014; Asfura-Heim and Espach, 2013; Miller, 2020; Rios, 2013; Wolff, 2020).

Hypothesis 1: As production increases, the conflict will increase.

In the case that participant B decided not to attack participant A, it could be because attacking A does not increase her own payoff or because both participants have settled an appropriate tax on production. In both cases, participant A could be better off as she does not need to allocate part of her resources to defend herself from attacks but to increase production. Her payoff could increase either due to an increase in the resources allocated to production or due to improvements in the production technology of production. In the case that production increases due to improvements in the technology of production, then participant B will increase her own payoff by attacking participant A. This leads to my second hypothesis:

Hypothesis 2: As production increases, the presence of cartels (participant B) will increase.

In the case that production increases due to an agreement between participants, then the technology of fighting changes and participant B will get more payoff than before using the same share of resources. In this case, participant B could increase her own payoff by devoting less resources to attacking participant A and using the extra for her own production. This leads me to my third hypothesis:

Hypothesis 3: As production increases and cartels (participant B) function as "stationary bandits," the conflict will attenuate (decrease).

Now assume, there is a third participant B' with the same set of actions (another cartel that, for instance, splintered from the first one) that was attracted by the increasing profits of participant A, and therefore, participant B' is also interested in getting the product of participant A. In the case of uncoordinated competitive theft between B and B', participant B' will either attack participant A in order to get their product or participant B in order to get territorial control. Both participants B and B' are interested in the monopolization of theft and protection of the taxgenerating subject, participant A. Given the presence of multiple thefts, the priority of participant B (and B') is to attack the other participant to gain control. In the face of multiple attackers, participant A has less incentives to produce, for instance, as she needs to devote more of her resources to defend herself. This leads to my third and fourth hypotheses:

Hypothesis 4: When there is presence of multiple cartels (participants B and B'), as production increases, the conflict against farmers (participant A) will decrease.

Hypothesis 5: When there is presence of multiple cartels (participants B and B'), as production increases, the conflict between cartels will increase.

1.3 Context

In this section, I present the context of the avocados and the agricultural sector in Mexico; the Mexican war on drugs and the increasing trend in violence since its start; and the potential interactions between these two activities.

1.3.1 Avocados and agricultural sector

Avocados are originally from the Americas but are currently produced on every continent. In 2019, the Americas produced 73.5% of the global production, and Mexico alone produced 2.3 million metric tons (MMT) making it the largest producer in the world.[2] Mexico is also the largest avocado exporter in the world, with nearly 0.95 MMT valued at over USD 2.4 billion exported to the United States (SAGARPA, 2017; USDA, 2020). Figure 1.1 shows the evolution of the volume of avocados produced in Mexico and the volume exported. In addition, it summarizes the number

[2]Data recovered from FAO stats. According to Miller (2020), Mexico, Peru, Chile, the Dominican Republic, and Colombia, account for 71% of the avocados grown in the world .

of municipalities producing the fruit. The agricultural sector – and avocado production– benefited from different trade agreements signed by Mexico starting in the 1990s. Beers, avocados, and berries are the three main products, valued from 2012 to 2018 at 3,768 USD million, 3,201 USD million, and 2,106 USD million respectively (SIAP, 2018). The latter two products are mainly exported to United States, and they are two of the products with a higher increase in the value exported. For instance, during the 1990s, the production of avocado was stable at around 0.8 million tons per year. After North American Free Trade Agreement (NAFTA) was in place, avocado production began to increase steadily reaching 2.3 million of tons in 2019. Avocado trade values increased by over 455% (USDA, 2020).

The United States banned the import of Mexican avocados in 1914 to protect the local production of the fruit claiming that the presence of pests in Mexico could damage local orchards (Miller, 2020). Starting in 1997, as part of NAFTA negotiations, avocados coming from the state of Michoacan were allowed to be imported to 19 states of the United States during the winter season. Mexican avocados were perfect to supplement avocado production in United States. Avocados produced in California have a peak season for harvesting in between April and July, whereas in Mexico they are produced all year long. In 2001, the United States government expanded the shipping zone and season for avocado imports, and since 2007, there are virtually no restrictions for avocados coming exclusively from the state of Michoacan. In 2016, the Animal and Plant Health Inspection Service (APHIS) lifted the restrictions placed upon avocados, allowing imports coming from any Mexican states[3] as long as they meet the requirements and were accompanied by a phytosanitary certificate (APHIS, 2016).

[3] Approvals for the import of avocados are made on a state-by-state basis. At the time of the extension rule, Jalisco was the first new Mexican state that met the requirements, apart from Michoacan, and it was authorized to export avocados to the United States in 2021. Subsequent Mexican states will be approved as each of them demonstrates its ability to meet the requirements (APHIS, 2016).

Although the majority of the Mexican avocado production is concentrated in the state of Michoacan, 29 out of 32 states in Mexico had at least one municipality growing avocado between 2003 and 2019. This production is used to meet the internal demand for avocados, and to export to other destinations, besides the United States, such as Canada and Japan. Each of these last two countries represents around 7% of the total exports of Mexican avocado and more than 90% of their respective imports of avocado. In the case of the United States, Mexican avocados represent a similar share of the total imports, although they are the largest buyers of Mexican avocados with around three quarters of the total export (Huang et al., 2023; SAGARPA, 2017; USDA, 2020). In addition, exports to the United States come almost exclusively from Michoacan, a state also known for the presence of criminal groups that have been recently linked with shortages of limes (Estancona and Tiscornia, 2022),[4] another crop that has increased its production to satisfy US demand.[5]

The lifting of the ban in the 1990s caused 'the green rush' in the state of Michoacan. Farmers and drug organizations replaced their production of marijuana with that of avocados, as the latter became a billion-dollar export business overnight (Miller, 2020). Criminal organizations 'taxed' farmers per hectare of orchards and, in some cases, they also 'taxed' government officials. Those farmers who opposed had their orchards burned and their families terrorized. Those who abandoned their orchards were forced to sign the land over to the criminal organizations. For in-

[4]For instance, for media coverage see https://www.nytimes.com/2022/05/04/world/americas/mexico-cartels-michoacan.html.

[5]Although avocados, limes, and berries have increased their demand in recent years, I focus on avocados because their production has increased consistently for more than 20 years, while the production of blueberries and raspberries increased after 2012. For instance, one of the factors driving the growth in recent years of the Mexican blueberries were subsidies (Wu et al., 2018). Additionally, among these agricultural products, avocados are the most important, while berries—like strawberries, blueberries, raspberries, and loganberries—are second. Lime is among the 20 more exported products in Mexico, valued at 513 USD million, one sixth of the value of avocados and about one fourth of that of berries. Moreover, the dollar value of avocado imports is higher than that of pineapples, bananas, and strawberries combined, despite of the United States importing more kilograms of all three of them.

stance, by the mid-2010s, *Los Caballeros Templarios* (The Knights Templars, a Drug Trafficking Organization based in Michoacan) was estimated to have direct ownership of about 10 per cent of the orchards in the state and indirect control of a much larger percentage (Miller, 2020).

1.3.2 The Mexican War on Drugs

For most of the 20th century, Mexico had a passive attitude toward the drug trade, with Mexican politics dominated by the Institutionalized Revolutionary Party (PRI) during this period. It was only the last PRI president, Ernesto Zedillo (1994-2000), and his successor, Vicente Fox (2000-2006), the opposition president from the right-wing National Action Party (PAN), that implemented security reforms and crackdowns (Chabat, 2010). The strategy mainly focused on destroying marijuana and opium crops in mountainous regions (Utar, 2018). Starting in December 2006, after the election of the PAN candidate Felipe Calderón, the Mexican government made fighting organized crime a major priority. During his second week in office, Calderón deployed 6,500 federal troops to combat trafficking, and by the end of his presidency, approximately 45,000 troops were involved (Dell, 2015).

After the initial deployment of troops in what was called *Operativo Conjunto Michoacán*, the escalating violence was unexpected and led to a bloody war between cartels to control drug routes and production of illegal substances (Salcedo, 2021). The crackdown of the drug cartels led to an impressive escalation of violence (Figure 1.2) due to confrontations between the state and the cartels, as well as between and within cartels (Battiston et al., 2022; Lessing, 2015; Rios, 2013; Shirk and Wallman, 2015). The war on drugs differently affected areas that had more presence of state armed forces (Atuesta and Ponce, 2017; Rios, 2013; Salcedo, 2021; Shirk and Wallman, 2015) and areas that supported the national government, i.e., were mayors where members of the same party as the president (Dell, 2015; Shirk and Wallman, 2015). Cartels began to expand to municipalities in an attempt to diversify their sources of income and to control any market that seems favorable, including markets for legal commodities (Estancona and Tiscornia, 2022; Piccato, 2022; Wainwright, 2016), in a multi-commodity criminal approach. For instance, the war on drugs pushed cartels into oil thefts, a new business that started to emerge shortly after (Battiston et al., 2022).

The intensification of the conflict led some criminal organizations to split themselves due to their leaders being captured or killed. For instance, after the death of the leader of the *La Familia Michoacana* cartel (The Michoacan Family), differences between successors led to the creation of the *Los Caballeros Templarios* cartel (The Knight Templars). Similarly, the foundation of the *Jalisco Nueva Generación* cartel (Jalisco New Generation) was due to a power vacuum and the association with a smaller faction of the *Milenio* cartel. These splints led to an increase in the number of municipalities with cartel presence and multiple cartels being active (Calderón et al., 2015; Rios, 2013; Sobrino, 2020). Figure 1.3 presents evidence of this expansion.

1.4 Empirical Strategy

1.4.1 Main specification

I am interested in whether the increasing production of avocados has led to changes in the intensity of conflict. The ideal experiment would compare the number of homicides in areas that have different avocado production levels for reasons otherwise unrelated to violence. However, I cannot consider such a comparison as it is reasonable to assume that the presence of avocado orchards could be affected by the conflict intensity. On the one hand, cartels may exert violence against avocado farmers who may decide to leave and stop producing, and on the other, farmers may decide to change their production of avocados to avoid being 'taxed' by cartels (see section 1.3).

To approximate such an experimental idea, I consider a Bartik-type or shift-share variable that combines geographical variation in avocado cultivation suitability and the growth in the demand for avocados in the United States, defined as follows:

$$Z_{mt} = \hat{S}_m \times \log(D_t)$$

where $\hat{S}_m$ is a predicted avocado suitability index, whose construction I explain in section 1.5.1.1, and it takes values between zero and one, where higher values means that municipality m is more suited to growing avocado. D_t is an external demand shock that proxies the growth of the avocado export sector.[6] This variable measures which areas are more suited to absorb the change in demand for avocados. Intuitively, this is a measure of *potential production* of avocados. To ease interpretation, I standardized this measure to have zero mean and unit variance. I discuss both variables in more detail below in section 1.5.

Using this measure of *potential production*, I estimate the following specification:

$$Y_{mt} = \beta Z_{mt} + \pi_m + \delta_t + \sum_{c \in W_m} [c \times \delta_t] + \varepsilon_{mt} \;, \tag{1.1}$$

where Y_{mt} is the homicide rate in municipality m and year t, which proxies for intensity of

[6] I am using the variable in logs rather than in levels to approximate the growth rate of the demand shock (Dube and Vargas, 2013; Mejía and Restrepo, 2015). Estimates in Table A.4 show that my results are not sensitive to the specification using the log of D_t. See appendix section A.4.1.2.

conflict. π_m and δ_t are municipality and year fixed effects, respectively. W_m is a vector of control variables interacted with year fixed effects. It includes the linear distance between the municipality and the state capital, and municipal characteristics observed in 2000, comprising the (log) population, the Gini coefficient, and a social gap index. The error term ε_{mt} is clustered at the municipality level to account for potential correlation across time within municipalities. To account for the sampling variance of the generated regressor $\hat{S}_m$, I block-bootstrap my standard errors preserving the municipality clusters. In addition to the block-bootstrap standard errors, to account for potential spatial correlation, I use a 500 kilometers radius to calculate the Conley (1999) standard errors. The parameter of interest β measures how changes in the areas more suited to absorb the demand for avocado (potential production) affect the homicide rate.

1.4.2 Identification

My identification comes from the differential suitability to cultivate avocados to evaluate how an external demand shock affects production in different municipalities. Specifically, the United States' demand for avocados is not correlated with changes over time in violence across municipalities in Mexico. Furthermore, municipalities that produce avocados either for internal consumption or to export to other countries should not be affected by this demand shock. Notice, I cannot leverage identification coming from the demand as I do not have multiple demand shocks (Goldsmith-Pinkham, Sorkin and Swift, 2020). In particular, I do not have data on other countries' consumption, such as Canada or Japan, the second and third most important buyers of Mexican avocados.

To achieve valid identification, the demand for avocados in the United States or the suitability to grow avocados should be exogenous to the intensity of conflict in Mexico. Notice, recent literature on shift-share specifications suggests that both variables do not need to be exogenous to have valid identification.[7] In my case, I argue that the demand for avocados in the United States is exogenous to cross-sectional conflict intensity in Mexico, as the latter has not interrupted the supply of avocados coming from Mexico. The growth in availability of avocados in the United States has come mostly from the import of Mexican avocados, although its share of total imports has been stable in the last decade. The cover the United States' excess demand for avocados, the volume of avocados imported from other countries –such as Peru or Colombia– has also been increasing.

On the other hand, I could also argue that suitability to grow avocados is exogenous to conflict as it is based on agroclimatic characteristics that should not affect conflict intensity. However, I construct the suitability index using observational data –i.e., whether or not a municipality has grown avocados– that could have been affected by the conflict, e.g., criminal organization forcing farmers to produce avocados (Miller, 2020). For this reason, I cannot claim that the avocado suitability index is completely exogenous.

A possible threat to my identification may be caused by municipal characteristics that may be correlated with the potential production, i.e., the suitability index (Ferri, 2022; Goldsmith-Pinkham, Sorkin and Swift, 2020; Jaeger, Ruist and Stuhler, 2018). For instance, areas more suited to growing avocado have higher levels of inequality or poverty, are less populated or far from an important market. To account for this potential threat, as a robustness to my main specification, I flexibly control for the interaction between these variables and year fixed effects in

[7] For instance, Borusyak, Hull and Jaravel (2021) show that their identification comes through exogenous shocks while allowing for endogenous shares while Goldsmith-Pinkham, Sorkin and Swift (2020) and Borusyak and Hull (2023) suggests that identification comes through the shares.

some of my specifications.

1.5 Data

My main data set consists of a panel of municipalities in Mexico for the period 2005-2019. Below, I describe the sources and information available for the main explanatory variable and the outcome variables. Tables 1.1 and 1.2 present the descriptive statistics for all the variables used in my analysis.

1.5.1 Measure of 'potential production'

1.5.1.1 Avocado suitability index

I exploit cross-sectional variation across municipalities using an avocado suitability index. Following Mejía and Restrepo (2015), I construct the index based on agroclimatic characteristics at the municipal level. To uncover the determinants of avocado production, I estimate the following model:

$$\begin{aligned} S_m = \quad & \alpha_0 + \alpha_1 altitude_m + \alpha_2 altitude_m^2 + \alpha_3 water_m + \alpha_4 water_m^2 \\ & + \alpha_5 soil_m + \alpha_6 soil_m^2 + \sum_{p=1}^{15} \alpha_p A_{mp} + \epsilon_m \end{aligned} \tag{1.2}$$

where $altitude_m$ is the average altitude of the municipality m, $water$ is a measure of water availability using the average annual precipitations (in mm) from 1990 to 2000, $soil$ is an index of soil suitability for common agricultural crops. To control for geological characteristics, A_{mp} is an indicator taking the value of one if a municipality m belong to Mexican physiographic region p. These regions do not necessarily overlap with administrative or political borders; therefore, a municipality can have part of its area in more than one of these regions, i.e., the dummies are not mutually exclusive (see section A.2 and Figure A.1). I consider a municipality to be part of physiographic region p if any part of the area of the municipality lie in the area of the region.

The outcome variable S_m takes the value of one if municipality m has reported a positive area cultivated with avocados in any year of the sample, using information from the Agricultural and Fisheries Information Service (SIAP) from the Mexican Secretariat of Agricultural and Rural Development (SAGARPA). This data contains municipal-level information about the number of hectares cultivated, the yield, and the annual volume and value of production for each municipality from 2003 to 2009 for a variety of crops, including avocados. Figure 1.1 summarizes this information by showing the total revenues from avocado production in current Mexican pesos over time. Alternatively, I use the information on the average productivity (yield) instead of the dummy indicator of avocado presence.

Table 1.3 presents estimates of variations of this model. In Column (1), I estimate equation 1.2 using a logit model and, in Column (2), a Linear Probability Model; in Column (3) and Column (4), I use the yield in levels and in logs using the inverse hyperbolic sine (I.H.S.) transformation, respectively, instead of the dummy indicator of avocado presence. My results suggest that the three variables present a positive but decreasing relationship with the presence (productivity) of avocados. This result is in line with other estimates that suggest that there are a wide range of values of water and altitude under which avocados can be cultivated, but they have optimal values within that range (De Haro, 2022; Miller, 2020). In terms of soil quality, the quadratic relation is less obviously interpretable as higher values means the municipality has a more favorable soil for different crops. A possible explanation is that more favorable areas are used to grow other cash crops. Overall, these agroclimatic variables explain 16.3 percent of the variation.

Using my estimates from Column (1), I create a measure of avocado suitability index S_m for all municipalities with agroclimatic data:

$$\hat{S}_m = f(altitude_m, water_m, soil_m; \hat{\boldsymbol{\alpha}}) \tag{1.3}$$

where $f(\cdot)$ is a flexible function of the estimates $\hat{\boldsymbol{\alpha}}$. In particular, I use a logit function to predict the index $\hat{S}_m$ so that it takes values between zero and one, where higher values represent areas more suitable for growing avocados. Alternatively, I use a linear function of the estimates when they come from a linear regression. Similarly, when I use yield as the outcome, the values of the index will fall outside the zero-one range and take positive values.

My index $\hat{S}_m$ varies across municipalities but not over time and its spatial distribution is shown in Figure 1.4. Panel a shows the areas that has avocado presence by summarizing the average annual production (yield), while panel b summarizes the avocado suitability index ($\hat{S}_m$) by quintiles. The areas more suited to grow avocado are clustered along the pacific coast (southwest of Mexico) and the Yucatan peninsula (east of Mexico) due to the water availability.

1.5.1.2 External demand

I exploit time variation in the demand for avocados using the availability of pounds per person for consumption in the United States. There is no data that allows me to recover the actual demand for avocados, therefore I proxy it with availability of the fruit. Figure 1.5 shows that availability in pounds per person increased fourfold in the last three decades, going from 2 pounds per person in 1990 to 8 pounds per person in 2020. This increase in availability comes mainly from imports as local production in the country has remained constant, or even decreased, in this period. In particular, the share of avocados that come from Mexico represents about 90% of the United States' total imports and 75% of the avocados consumed in the country. To keep with the demand, United States avocado imports have increased from other countries such as Peru and, more recently, Colombia (Figure 1.6).

1.5.2 Homicide and conflict data

Data on homicides come from the National Department of Health Information (SINAIS). It consists of individual level information for which a death certificate was generated and includes detailed information on the cause of death (including homicides), the location of occurrence, and individual characteristics such as age, gender, and occupation. Using the location of residence and cause of death, I calculate the homicide rate per 100,000 people in each municipality-year.[8] Based on the victim's occupation, sex, and age reported, I classify the victim into one of three groups and calculate their respective homicide rate. Figure 1.7 summarizes how I classify homicides. The first group consists of the victims whose occupation was agricultural workers. The second group consists of male victims between the ages of 15 and 40 whose occupation was not agricultural workers, as this population group is more likely to take part in cartel activities.[9] The third group consists of any victims that do not belong to either of the first two groups.

[8]The correlation between the victims' location of residence and location of occurrence is 0.9902. See section A.3 for details on the classification of homicides.

[9]There is no other information that allow me to classify the victims as cartel related. Other occupations include professionals, merchants, personal and security services, artisan workers, heavy machinery and drivers. To create mutually exclusive groups, I am excluding agricultural workers from the group of (possibly) cartel related. I explore the sensitivity of this classification in section 1.7 and section A.4.

I focus on homicides and use the SINAIS data for three reasons. First, homicides are the most reliably registered crime, although the kidnapping, extortion, and disappearances and mass graves are also associated with Mexican crime groups (Prieto-Curiel, Campedelli and Hope, 2023; Shirk and Wallman, 2015).[10] Second, the SINAIS data includes detailed information on the occupation, sex, and age of the victims, which allows me to improve with respect to previous studies. Third, SINAIS data is more accurate than other sources, such as police reports, as it contains information about all individuals for whom a death certificate was generated. Police report data suffers from underreporting as file cases are not open and investigated (Atuesta, Siordia and Lajous, 2018).

1.5.2.1 Presence of cartels

Data on the presence of cartels comes from the Mapping Criminal Organizations project.[11] This is the most complete attempt to map criminal organization in Mexico at the cartel-municipal-level from 1990 to 2020. It approximates the presence of 75 different cartels in a municipality by analyzing whether the cartel is mentioned in local news (Sobrino, 2020).[12] Using this data, I identify whether a cartel is present (extensive margin) in each municipality-year, and how many cartels there are (intensive margin). Furthermore, to identify the cartels with likely greater national presence, I split cartels into traditional –the 9 main cartels that have been present since the

[10]I use information from the National System of Public Security (SPP) on the number of homicides, extortion, kidnapping, and threat cases reported to the police in section 1.8 below. I do not have information on disappearances that allow me to identify possible under-report of homicides, as there is not death certificate generated in those cases.

[11]I thank Fernanda Sobrino who shared this data with me.

[12]This data can suffer from measurement error as actual presence could be related to no news reporting (Atuesta, Siordia and Lajous, 2018); however, this is the more complete data set at this level. Coscia and Rios (2012) report the present during the 1900s and the 2000s. Other sources report the presence of more criminal organizations at the national level, but they do not cover a prolonged period. For instance, Prieto-Curiel, Campedelli and Hope (2023) identifies 150 different organizations in 2020 based on the *Drug Policy Program* (PPD) data from the *Centro de Investigación y Docencia* (CIDE).

1990s– and non-traditional cartels –the newly formed cartels or splinters of traditional cartels.[13]

1.6 Results

My main results for the homicide rates, using the classification of the victims, are presented in Table 1.4. Column headers report the outcome variable and rates are expressed in 100,000 people using the municipality population of 2010. Estimates in all columns include municipality and year fixed effects, and even columns include the interaction between municipal-level controls and year fixed effects. To ease interpretations, the mean of the dependent variable in the sample is reported in the bottom of the table.

1.6.1 Main results: homicide rate

I find a positive effect of increasing demand for avocado on the homicide rate. I also find that the increase in the homicide rate is driven by an increase in the homicide rate of agricultural workers.

Focusing on the homicide rate of agricultural workers in Panel A and the most demanding specification (Column (4)), a one-standard-deviation increase in the area more suited to absorb the demand for avocados increases the homicide rate of agricultural workers by 1.06 additional homicides per 100,000 people. This effect is statistically and economically significant and represents a substantial size effect of 18.3% with respect to the 5.76 mean homicide rate of agricultural workers. In general terms, this additional death represents a 6% increase with respect to the 16.96

[13] The Mexican prosecutor's office and the DEA recognized the Gulf Cartel, the Juarez Cartel, *La Familia Michoacana*, the Tijuana Cartel, the Sinaloa Cartel, Beltran-Leyva Organization, Jalisco New Generation Cartel, Los Zetas, and The Knights Templar as the 9 main Drug Trafficking Organizations (De Haro, 2022; Sobrino, 2020). For a brief historical description of cartels formation in Mexico, see Atuesta and Pérez-Dávila (2017).

mean homicide rate in the municipality. I do not find an effect statistically different from zero for the homicide rate of (potentially) cartel related individuals nor for the rest of the population.

To account for potential spatial correlation between municipalities, I report the Conley (1999) standard errors using a 500 kilometers radius, following Sobrino (2020), in addition to the block-bootstrapped standard errors in Table 1.4. The former are more conservative than the block-bootstrapped standard errors by around 10% while keeping their significance at the conventional levels. Given the similarity of the standard errors, I only report the block-bootstrap standard errors at the municipal level in the rest of the tables.

A possible threat to my identification may be caused by differences in municipal characteristics that may be correlated with my explanatory variable, as discussed above (Goldsmith-Pinkham, Sorkin and Swift, 2020). To account for this possibility, even column estimates in Table 1.4 include the interaction between municipal-level control and year fixed effects. Reassuringly, the magnitude of the estimated coefficients is robust to the inclusion of controls.

1.6.2 Avocado production

Using information on actual avocado production, I could estimate its relationship with the intensity of conflict using a two stage least squares. Specifically, I estimate:

$$y_{mt} = \theta avocado_{mt} + \tau_m + \psi_t + \sum_{c \in W_m} [c \times \psi_t] + \eta_{mt} \,,$$

where $avocado_{mt}$ is a measure of avocado production. τ_m and ψ_t are municipality and year fixed effects, respectively. Municipal characteristics W_m are as defined above and are interacted with year fixed effects. The parameter of interest is θ. The first-stage is as follows:

$$avocado_{mt} = \gamma Z_{mt} + \lambda_m + \phi_t + \sum_{c \in W_m} [c \times \phi_t] + \upsilon_{mt} \ ,$$

where λ_m and ϕ_t represent municipality and year fixed effects.

I show that my measure of potential production explains the production of avocado. Table 1.5 reports the first-stage estimates using different measures of avocado production in levels and in logs.[14] Specifically, I use four separate treatment variables: the total area cultivated (in hectares) with avocado, the volume and the value of the avocado production, and the value of the production per capita. I find a positive relationship between my measure of areas more suited to absorb the demand for avocados and the actual production of avocados, both in levels and in logs. Focusing on the main specification, which is in Column (8), a one-standard-deviation increase in area more suited to absorb the demand for avocado increases the value of avocado production per capita by 1,319 Mexican pesos.

I find a positive effect of avocado production on the conflict intensity using potential production as the excluded instrument. For intuition purposes, I use the value of avocado production per capita in the municipality as my measure of avocado production. Focusing on Column (4) in Panel B of Table 1.4, I find that an increase in 1,000 Mexican pesos per capita, increases the homicide rate of agricultural workers by 0.76 additional deaths per 100,000 people. Remarkably, this estimate is similar in size to the reduce form estimate reported in Panel A.

I take with caution the results coming from this specification as the monotonicity assumption may not be satisfied. Assuming that the measure of potential production is exogenous to the conflict intensity, as discussed previously (see section 1.4.2), my instrument of potential pro-

[14]I use the inverse hyperbolic sine instead of the natural logarithm as there is presence of zeros in my data. I show that the results are not driven by this transformation in section A.5.

duction is relevant and meets the exclusion restriction. However, I cannot guarantee that the monotonicity assumption holds. In particular, even if I can assume that municipalities with low potential to produce avocados are producing them (compliers) due to the increasing demand, I cannot completely rule out that municipalities with high potential to produce avocados are not producing them (defiers).

1.6.3 Time-varying unobserved heterogeneity

One possible threat to my identification strategy is the presence of time-varying unobserved heterogeneity that is not captured by the fixed effects. The presence of this type of heterogeneity is more likely to happen with large longitudinal data and it will bias my estimates (Millimet and Bellemare, 2023). To account for this possibility, I re-estimate my main specification using the First Difference (FD), Twice-First Difference (TFD), the Rolling First Difference (RFD), and the Rolling Twice-First Difference (RTFD) estimators.[15] I report in Table 1.6 these alternative estimates, in addition to the fixed effects (FE) estimates to ease comparison.

My results suggest that there is presence of unobserved characteristics that are not accounted for by the fixed effects estimator. For instance, the various policies adopted around the Mexican "war on drugs" could have generated structural breaks in the homicide rate that are not accounted for the fixed effects. Indeed, the FD estimates on the overall homicide rate is larger than those of the FE. This increase is largely explained by an increase in the homicide rate of the rest of the population, which speaks about the changes in the conflict intensity in Mexico.

[15]The Rolling First-Difference model and the Rolling Twice First-Difference model can be estimated with or without a constant term. The First-Difference model with a constant assumes that $E[\alpha_i^{t',T} - \alpha_i^{t',s}] = c$, where $\alpha_i^{s,s'}$ is the time-varying part of the unobserved heterogeneity over period defined by s and s', and t' is the period where there is a structural break, since $E[\alpha_i]$ is not assumed to be zero (Millimet and Bellemare, 2023).

Similarly, the FD and TFD estimates on the homicide rate of agricultural workers are larger than those of the FE, while they still explain the change in the overall homicide rate. Therefore, I can assume that the FE estimates represent a lower bound of the true effect.

1.7 Robustness

There might be concerns about the way the explanatory and outcome variables are constructed, or that the units in my sample drives the results. To alleviate these possible concerns, in this section I present different robustness tests for the main estimated effects. I follow Eggers, Tuñón and Dafoe's (2023) typology of placebo tests and split my analysis into three parts: alternative (i) explanatory variables, (ii) outcome variables, and (iii) populations to assess my estimates effects.

1.7.1 Alternative explanatory variables

One alternative explanation to the increase in the homicide rate of agricultural workers is that this population group is more targeted, even if they are not producing avocados. To rule this possibility out, I perform placebo tests for my explanatory variable using alternative crops that should not affect the intensity of conflict. In addition, I also use alternative measures of avocado suitability and external demand to show that the increase is driven by changes in the demand for avocados.

1.7.1.1 Suitability for alternative crops

I show that the increase in the homicide rate of agricultural workers is experienced in areas which are suitable for growing avocados, while areas which are suitable for growing other crops are not experiencing similar increases in homicides. To do this, I calculate a suitability index for growing different crops and interact it with the demand for avocados. Specifically, I look at the suitability for growing different fruit trees, namely apples, peaches, pears, oranges, and lemons. My results suggest that the increase in the homicide rate of agricultural workers is not related to the suitability of growing other crops, but it is explained by the suitability for growing avocados (Figure 1.8). This result goes in line with the profitability of avocado production and its high-value export (see section 1.3). The only exception is for the areas more suited for growing lemons, which can be explained by the presence of both crops in areas with cartel presence (Miller, 2020).[16]

1.7.1.2 Alternative shift-share measures

It might be the case that my proposed measure of potential production and its functional form are driving my results. To address these concerns, I redefine my measure by modifying either the share part –alternative measures of suitability for growing avocado– or the shift part of the variable –alternative exogenous demand shocks. Specifically, I use the different measures of avocado suitability reported in Table 1.3 (see section 1.5.1.1) and the share of agricultural area devoted to avocado production in 2003, the first year I observe information at the municipal level. For the alternative shocks, I use the growth of the avocado industry in Mexico, the global

[16] For media coverage, for instance see https://www.npr.org/2022/02/19/1081948884/mexican-drug-cartels-are-getting-into-the-avocado-and-lime-business.

production excluding Mexico, and the number of searches on Google trends. My results are robust to using alternative shares, sources of shocks, and to not taking the log of the demand shock. My estimates are robust in terms of sign and significance, although they present different point estimates coming from differences in the measure units (Tables A.3 to A.5 in section A.4). The correlation between the demand for avocados in the United States and the different demand shocks is above .94, except with the Google search that is .80 which could explain the loss in significance.

1.7.2 Alternative outcome variables

To investigate how sensible my results are to the classification of the victim of homicides, I use the criteria of occupation, and age and gender to investigate potential overlaps between the victims' profiles. In addition, using the residence of the victims, I calculate the main outcome variables for those that lived in rural areas.

1.7.2.1 Alternative classification of homicide victims

To show that my results are not sensible to the classification of victims, I separately estimate effects by occupation, and age and gender (see Figure 1.7). Specifically, I separate the homicide rate for agricultural workers into males between the ages of 15 and 40 –the population group most likely to take part in cartel activities– and the rest of the agricultural workers. Similarly, I compute the homicide rate of all male individuals between the ages 15 and 40, regardless of their occupation. I estimate equation 1.1 using the homicide rate of these three groups in Table 1.7. To ease comparisons, I report in Column (1) and (2) the estimates for the agricultural workers;

and in (9) and (10), the (potentially) cartel related. My results suggests that the increase in homicides of agricultural workers comes in particular from male agricultural workers between the ages of 15 and 40 (Column (4)). This latter group explains the increase in the homicide rate of agricultural workers (Column (2)) and, also, in the homicide rate of males between the ages 15 and 40 (Column (8)). Homicides of agricultural workers outside the age range (Column (6)) are unaffected by avocado demand nor are the homicide rates of individuals in the age range who are not agricultural workers (Column (10)). This result is not surprising as other studies have linked agricultural workers to criminal activities through the opportunity cost channel (Dube and Vargas, 2013; Guardado and Pennings, 2020). Another possibility is the rise of vigilantism in Mexico that has created self-defense groups of agricultural workers to protect local production from criminal organizations (Althaus and Dudley, 2014; Wolff, 2020).[17]

1.7.2.2 Homicides for residents of rural areas

Using the recorded residence of the victims, I restrict my sample to homicide cases with victims living in rural areas as they are more likely to engage in agricultural activities. For instance, about half of the victims of homicide in a municipality lived in rural areas, whereas for agricultural workers, it was the majority of the victims. My results are robust to this classification and are reported in Table 1.8. In particular, the increase in the homicide rate of individuals in rural areas is solely explained by the increase in the homicide rate of agricultural workers whose place of residence was the rural area. Although the point estimate is smaller than that without this restriction, the size effect is around 16% when compared to the sample mean.

[17] In section A.4, I presents estimates for the main occupations reported and show that agricultural workers explain the changes in the overall homicide rate.

1.7.3 Alternative populations

To rule out that my results are driven by outliers or clusters of municipalities, I exclude states, large cities measured by population, and extreme values in the reported homicides.

1.7.3.1 Excluding states

To show that my results are not driven by a particular cluster of municipalities, I run my estimates excluding from my sample each one of the 32 Mexican states at a time.[18] It may be the case that most of my variation in the suitability for growing avocados come from municipalities that have a higher intensity of conflict. For instance, this could be the case of Michoacan, the largest producer state of avocados in Mexico, which is also notable for the intensity of conflict (Althaus and Dudley, 2014; De Haro, 2022; Wolff, 2020). I rule out this possibility and show that my results are robust to exclude every one of the 32 states. Figure 1.9 presents the estimates for the homicide rate in agricultural workers in panel a and for (possibly) cartel related in panel b. Surprisingly, a notable exception is the exclusion of Jalisco from my estimates, which increases the homicide rate of (possibly) cartel related individuals. I would have expected, if something, an opposite effect to the exclusion of Jalisco as it is both a producer of avocados and presents high conflict intensity. A possible explanation is cooperation between criminal organizations that reduces conflict intensity. I explore the role of cooperation in section 1.8.

[18] Mexico is comprised of 32 federal entities: 31 states and the capital, Mexico City. States are the first level administrative division, which is further divided into municipalities. For simplicity, I refer to 32 Mexican states.

1.7.3.2 Excluding large cities

I show that my results are not driven by municipalities that have high homicide rates. Mexico's conflict has claimed as many lives as war zones, including cities with homicide rates above 200 deaths per 100,000 people (Lessing, 2015; Rios, 2013). However, these violent cities are not small but rather big and mid-size cities. For instance, according to Rios (2013), the population in the top 5 most violent municipalities ranges from around 650 thousand (Acapulco) to just below 2 million inhabitants (Tijuana, the second largest city by population in Mexico). In Figure 1.10, I report my estimates restricting the sample to municipalities below different population thresholds and show that my estimates are not driven by the homicide rates of big cities.[19] However, note that for municipalities below 100 thousand inhabitants, the estimated effect doubles in size. Interestingly, the estimates on homicide rate of (possibly) cartel related individuals become more negative as the population decreases.

1.7.3.3 Outliers in homicides

I rule out the possibility of outliers in the number of homicides by dropping from my sample those observations whose number of homicides is in the top 1 percentile. Homicide information may be biased as the location of occurrence is not necessarily the same as the location where it is registered. For instance, the average homicide rate coming from police reports is 26.2 (Table 1.1), while the homicide rate for the same period using information from death certificates is 19.5 deaths per 100,000 people. If the latter is related to the suitability for growing avocados or to conflict intensity, then they will be biasing my estimates. I exclude the observations where the

[19] Table A.7 in appendix further explores this exclusion.

number of homicides (in levels) is above the 99th percentile for the respective homicide rate in the outcome variable. I show that my estimates are robust in size and significance to this exclusion in Table A.8 in the appendix.

1.8 Mechanisms

My results so far suggest that the demand for avocados increases the homicides of agricultural workers in areas more suited to absorb that demand for avocados. These findings are in line with other studies recently carried out with different agricultural commodities responding to production shocks, e.g., changes in prices (Cisneros, Hellmundt and Kis-Katos, 2022; De Haro, 2022; Dube and Vargas, 2013; Estancona and Tiscornia, 2022). In this section, I explore the mechanisms plausibly driving the increase in the number of homicides as a response to the increasing demand for avocados. First, I study other forms of criminal activities and the role of cartels. Second, I empirically test the mechanisms of cooperation between producers and cartels, and competition between cartels. Finally, I study whether crop production changes living standards (Boone, Kaila and Sahn, 2022).

1.8.1 Conflict and criminal organizations

1.8.1.1 Criminal activities

Changes in the demand for avocados could affect other types of criminal activities, beyond homicides, such as kidnappings, threats, or extortion (Buonanno et al., 2015; Shirk and Wallman, 2015; Wolff, 2020). I explore whether the increase in homicides is accompanied by changes in these other activities. In this section, instead of using information from death certificates, I use data from police reports including data on homicides as a robustness check to the source of information. Table 1.9 reports the estimates for the criminal activities coming from police reports. I find that a one-standard-deviation increase in the areas more suited to absorb the demand for avocados increases the homicide rate by 5 additional deaths. Although different in magnitude, the size effect is also around 20 percent, similar to the effect using information from death certificates, as the mean homicide rate is 26 deaths per 100,000 people.[20]

When I turn my attention to other types of criminal activities, I find that a one-standard-deviation increase in the areas more suited to absorb the demand for avocados decreases the extortion rate in 1.64 cases per 100,000 people, while the same change does not affect the threat and kidnap rates per 100,000 people. These are surprising results as I would expect these three measures to have a positive point estimate. A possible explanation for these results, particularly in the case of extortion, is that individuals are less willing to file a report. This can be the case when cartels have a preference for killing over extorting, or when there exists a compromise between the two sides to overcome the *commitment problem* (Blattman, 2022), i.e., farmers and cartels agree on a rate of taxation acceptable to both parties. This would also explain the null effect on threats and kidnappings. Alternatively, cartels may decide to act–i.e., killing– if farmers are unwilling to pay the extortion. This will bring the homicide rate up and the extortion rate down. However, since I do not observe detailed information on who is being extorted, it is not possible to validate, or rule out, these explanations. I further explore the role of cooperation in section 1.8.2.

[20] I explore differences in the sources of information in section A.3.

1.8.1.2 Presence of criminal organizations

I explore whether changes in the homicide rate are explained by changes in the presence of criminal organizations, or cartels. An unexpected consequence of the Mexican war on drugs was the proliferation and expansion of cartels throughout the country (Atuesta and Ponce, 2017; Atuesta and Pérez-Dávila, 2017; Calderón et al., 2015; Rios, 2013; Shirk and Wallman, 2015; Sobrino, 2020). Using data on cartel location coming from the Mapping Criminal Organizations project (Sobrino, 2020), I re-estimate equation 1.1 using separately both the presence of cartels as dependent variable, and the total number of cartels. I do this to capture extensive and intensive margins, respectively. In addition, I differentiate cartels between traditional –the 9 main Mexican cartels that have been present since the 1990s– and non-traditional cartels –newly formed cartels or splinters of traditional cartels (see section 1.5).

Table 1.10 presents my estimates for the extensive and the intensive margin, including those places with zero cartels. My results suggest that a one-standard-deviation increase in the areas more suited to absorb avocados increases the probability of having a cartel present, both traditional and non-traditional, and it also increases the number of cartels present. Although I find a positive effect on the extensive margin, these are similar to those of intensive margin as the average number of cartels is less than one. I further explore the extensive margin in section 1.8.2.2.

To study whether cartels mediate the increasing homicide rate, I re-estimate equation 1.1, interacting my measure of potential production with an indicator variable of cartel presence in the municipality in the previous period, i.e., heterogeneous effects by "contemporaneous" cartel presence. Table 1.11 presents the estimates using the classification of cartels for the homicide rate of agricultural workers in the first 6 columns; while in the last 6 columns, for the homicide rate of (possibly) cartel related individuals. I relegate to the appendix the table with the estimates for the other two main outcomes (See Table A.9). My results suggest that the increase in the homicide rate of agricultural workers is not driven by the presence of criminal organizations. This is unexpected as I would have imagined that the presence of cartels would lead to more homicides. On the other hand, the results for the homicide rate of (possibly) cartel related individuals present a different picture. The presence of traditional cartels decreases their homicide rate, while the presence of non-traditional cartels increases their homicide rate. A possible explanation for these results rely on cooperation and competition, which I explore in the next section.

1.8.2 The role of cooperation and competition

1.8.2.1 Historical cartel presence

As a complementary analysis, I study whether there is a differential response in areas with prior presence of cartels, as a proxy for cooperation between farmers and criminal organizations. The lack of trust in government is important to understand the role of cooperation and vigilantism, as government officials were involved with drug trafficking organizations at the local level (Shirk and Wallman, 2015; Wolff, 2020). These criminal organizations acted like "stationary bandits" by charging locals for protection (Olson, 1993) and fighting the state not to conquer it but to constrain it (Lessing, 2015). Given that the presence of "contemporaneous" cartels might be due to the increasing demand for avocados, I use the "historical" presence of criminal organizations to explore the role of cooperation and report my results in Table 1.12. In particular, I assume that a municipality had cartel presence if there was at least one cartel in any year of the decade of 1990s.[21] An advantage of using historical rather than contemporaneous cartel presence is that, newly arriving criminal organizations might focus on rent-seeking rather than focusing on territorial control, where cooperation is more likely to happen.[22] Column (4) estimates suggest that presence of historical cartels fully attenuates the increase in the homicide rate of agricultural workers; however, it does not have an effect on the homicide rate of (possibly) cartel related individuals. This result is similar to that of Crost and Felter (2019), who find that an increase in the price of bananas leads to a decrease in violent incidents in insurgency areas within the Philippines.

1.8.2.2 Competition between criminal organizations

In this section I explore whether cartel competition explains the increase in the homicide rate using information about presence of multiple cartels, as a proxy for competition.[23] The proliferation of cartels led to vigilantism and fights for territorial control between criminal organizations to establish a monopoly of looting (Blattman, 2022; Lessing, 2015; Olson, 1993). In particular, other research has shown that newly formed cartels have relatively limited influence over drug markets and routes, therefore they battle to take control over a "profitable" territory

[21] Note, before 2007 there were only traditional cartels (see Figure 1.3). Data on cartel presence from the Mapping Criminal Organizations project relies on text analysis of newspapers, which could be less accurate for the decade of 1990s.

[22] There are different explanations for cooperation. Blattman (2022) suggests that societies have managed peaceful competition when there is some *interdependence* between farmers and criminal organizations, or when they create some *rules* or social norms that both follow. Alternatively, Lessing (2020) suggests that the reduction in violence is associated with non-state actors exerting some way of governance over local communities. Indeed, in some instances, locals preferred crime organization leaders from the area to *foreign* crime organizations (Asfura-Heim and Espach, 2013; Wolff, 2020), as this allowed them to agree on a rate of taxation that was acceptable to both parties, even if this implied cartels lowering their 'taxes' as a mean to avoid confrontation (Miller, 2020).

[23] The presence of multiple cartels does not rule out cooperation between them. For instance, cartels allied to confront a common enemy, to obtain protection, or to strengthen debilitated groups. The effect of this cooperation on conflict varies with the conditions under which the cooperation takes place. For instance, there is an increase in the number of violent events when an emancipated faction join forces with a third cartel to clash with their former boss. However, there is larger increase in the number of violent events when cartels fight one another, specially following a fragmentation (Atuesta and Pérez-Dávila, 2017).

(Atuesta and Pérez-Dávila, 2017; Battiston et al., 2022; Estancona and Tiscornia, 2022; Rios, 2013; Shirk and Wallman, 2015; Wainwright, 2016). For instance, Prieto-Curiel, Campedelli and Hope (2023) estimate that more than half of the country's casualties result from fights between the smallest and the largest cartels; hence, the importance of differentiating the largest cartels –or traditional– and the smallest cartels. Using information on the number of cartels present in every municipality-year, I create an indicator variable for municipalities with only one cartel and another for municipalities with two or more, separately for traditional and non-traditional cartels, as a proxy for competition.[24] I interact these indicator variables with my measure of potential production and report my estimates in Table 1.13, focusing on the homicide rate of agricultural workers and the homicide rate of (possibly) cartel related individuals. Similar to cooperation, my results suggest that cartel presence in areas more suited to absorb the demand for avocados does not explain the increase in the homicide rate of agricultural workers. However, the presence of at least one non-traditional cartel is associated with an increase in the homicide rate of (possibly) cartel related individuals. Notably, the largest increase in violence happens when there is presence of two or more non-traditional cartels in areas more suited to absorb the demand for avocados, whereas there is no association when there is presence only of traditional cartels. This result is consistent with the expectation of small cartels competing over resources and territorial control.

[24] I simplify my analysis by focusing on 2 or more cartels rather than multiple categories for municipalities with presence of multiple cartels. This simplification assume that the presence of two groups has the same effect as the presence of, say, 10 groups. Atuesta and Ponce (2017) suggest that the violence between cartels present an inverted-u shape with respect to the number of criminal organizations due to fragmentation. However, this result only holds for confrontations and violence between groups and not necessarily against agricultural workers.

1.8.3 Economic benefits

1.8.3.1 Labor market and structural transformation

Using data from the National Employment Survey (ENOE), I analyze whether the demand for avocados brought changes in the labor market.[25] Previous studies have suggested that production booms could bring changes in the labor market, in particular that of export crops (Boone, Kaila and Sahn, 2022; Crost and Felter, 2019). These changes can come either through more employment or through more income (from wages). I use data on these variables and show that there are no changes in workers income, suggesting that the change in demand is not attracting more workers via higher wages (Table 1.14). In other words, profits from avocado production are not passed to workers. On the other hand, in terms of employment, I find that the areas more suited to absorb the demand for avocado experience an increase in the unemployment rate. Although this is not expected, a possible explanation comes from the nature of the survey that under-represent workers in rural areas.

An alternative explanation could be that the municipalities are more broadly undergoing a structural transformation towards the agricultural sector.[26] To test this hypothesis, I calculate the share of workers in the main sectors, namely agriculture, construction, manufacturing, services, and commerce. My results suggest that areas more suited to absorb the demand for avocado experienced a reduction in the share of workers in the manufacturing sector (Table 1.15). In

[25]The use of this survey have two caveats. First, the ENOE is not representative at the municipality level but rather at the state level and for 39 cities, underrepresenting the rural area. Second, the survey is collected quarterly, but I am using yearly information. To do so, I average the four quarters of the year. By doing this, I may be covering some of the heterogeneity including the seasonality of crops. However, given that the avocados can be harvested and exported all year long, I am not that concerned on missing seasonal effects.

[26]For instance, Uribe-Castro (2021) suggests that the development of the coffee sector in Colombia reduced the supply of skill workers and slowed down structural transformation, affecting the manufacturing activity in the long run.

addition, the agricultural sector estimates are positive, although they are not significantly different from zero. These results need to be taken with caution since the Mexican war on drugs displaces skill intensive jobs (Aldeco, Jurado and Ramírez-Álvarez, 2022; Salcedo, 2021). Therefore, areas where there is an increase in the homicide rate also present a recomposition in terms of the labor force.

1.8.3.2 State capacity

The increase in demand for avocados and for avocado production should be associated with an institutional response in the form of tax collection and or public spending. Previous studies have linked economic booms with more state capacity, measured as tax collection, or more public investment (e.g., Boone, Kaila and Sahn, 2022). Using fiscal reports on tax collection and expenses, I calculate the share of revenues coming from taxes and from production taxes, and the share of taxes from the latter to capture tax collection, as well as the share of expenses devoted to public investment and to personal expenses. I test this hypothesis by estimating equation 1.1 using these outcomes and report my results in Table 1.16. I find that an increase in the area more suited to absorb the demand for avocado is not associated with changes in public expenses; however, the same increase is associated with more revenues from taxes, particularly from production taxes. Notice, due to data availability, I do not have information about revenues from production taxes for all the municipalities in my sample. Hence, I take this last set of results as suggestive evidence for criminal going after the increasing revenues of the avocado business rather than going after farmers.

1.9 Conclusion

I study the increase in avocado production and its unintended consequences on conflict intensity in Mexico. My empirical strategy estimates the impact of avocado production exploiting a combination of cross-sectional variation from a suitability index for growing avocados, and time variation coming from the increased demand for the fruit in the United States. My analysis goes beyond estimating the effects of a resource boom on conflict intensity. Mexico hosts a large number of criminal organizations that fight over the monopoly of looting and territorial control. I not only estimate the impact of the resource boom on violence but additionally I study heterogeneous responses by cartel presence; that is, whether competition between criminal organizations and cooperation between criminal organizations and local farmers drives the changes in conflict.

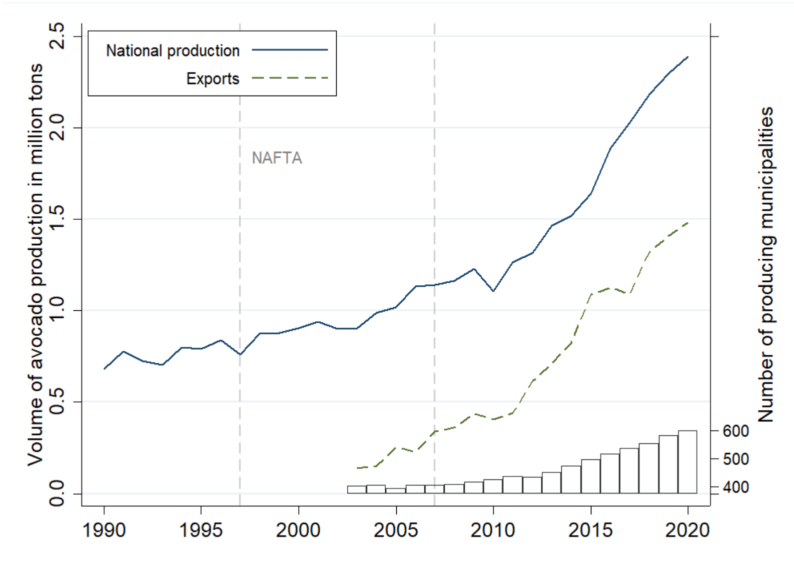
I first show strong evidence of a positive relation between the areas more suited to absorb the demand for avocados and the homicide rate. The effect is strongest on deaths of agricultural workers. I show that contrary to standard beliefs, cartels are not driving this change in the homicide rate. I find that historical cartel presence, which proxies for the degree of cooperation between cartels and farmers, fully attenuates the increase in the homicide rate of agricultural workers; on the other hand, presence of two or more cartels –competition– shifts homicides from agricultural workers to individuals who are involved with cartel activities. This result provides suggestive evidence in favor of the hypothesis that cartels fight over profitable territories.

My results shed light on the role of criminal organization to explain the benefits of increasing production of a high-value export crop. However, they should be taken with caution, as I cannot suggest that allowing criminal organizations to rule will reduce the number of homicides. Moreover, the increasing number of small cartels seems to be causing the majority of deaths.

Notice, I show contemporary effects of the effects of avocado production on conflict. Further research is needed to better understand whether in the long-run production of high-value export crops, such as avocados, could benefit the local economy and whether it is possible to terminate with small cartels, for instance, by adopting the production of avocados.

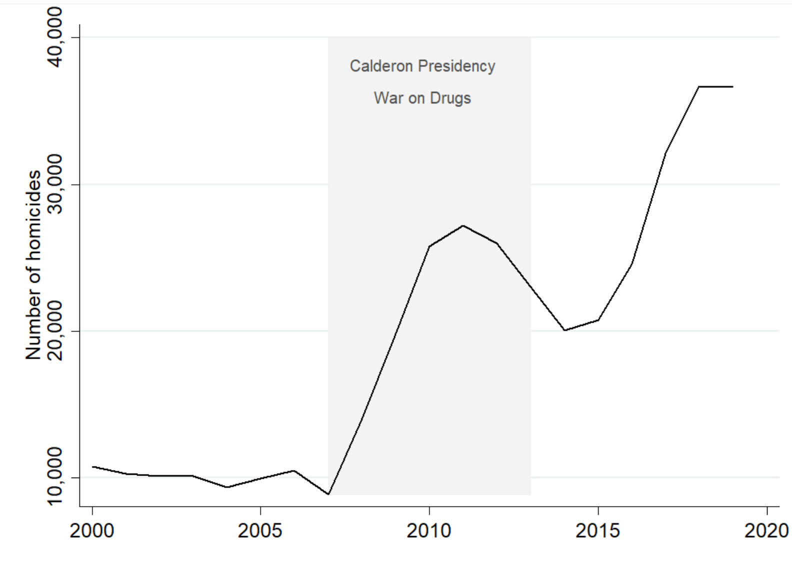
The extent of my analysis has some limitations. First, even though I find a positive relationship between cartel presence and avocado production, I cannot discern the exact nature of cartels' direct involvement in avocado production; yet there is anecdotal evidence of this phenomenon (Miller, 2020; Piccato, 2022). Second, this study does not address the possibility of cartels using avocado production as a means to smuggle drugs into the United States nor does it address the possibility of using avocados to launder money. For instance, Michoacan is Mexico's most important producer of avocados, its second most important producer of limes, and it is additionally heavily involved in the production and trafficking of illegal drugs. It has a unique location along the Pacific Coast including the port of Lazaro, which is the most important commercial gateway with Asia (Althaus and Dudley, 2014). Limited data availability prevents me from exploring those avenues of cartel interference in avocado production. Finally, because my analysis is limited to the effects of avocado production on conflict; I also cannot speak to the environmental effects of avocado production. Further research is needed to determine the impact of avocado production on land use as deforestation has been linked with both avocado production and criminal organization activity.

Figure 1.1: Avocado production and NAFTA



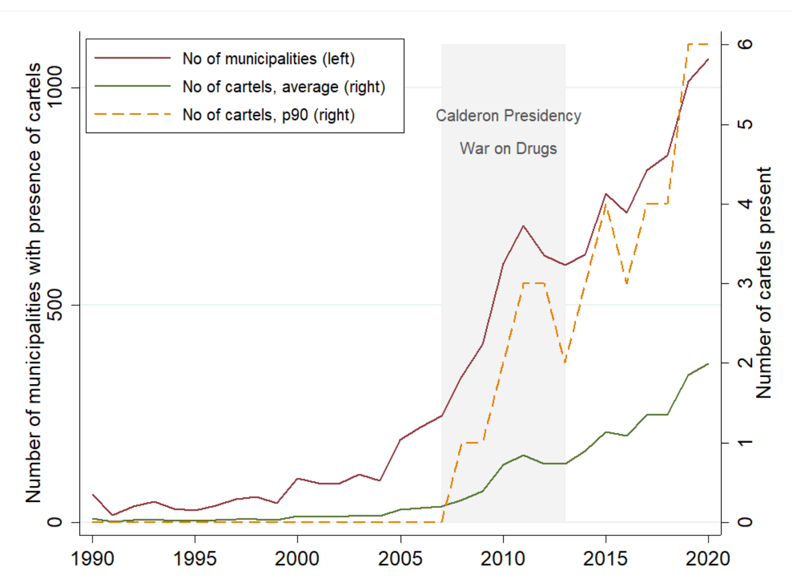
Note: Own calculation based on information from SIAP and SIAVI. Total production of avocado and exports are in million tons. Dashed vertical lines represent changes in NAFTA regulation. Avocados were first exported in 1997 under restriction of destiny and season. In 2007, restrictions are lifted and avocados are allowed to be exported to all the states and all year long in United States. Bars summarizes the number of municipalities producing avocados (right axis). Municipal level data starts in 2003.

Figure 1.2: The War on Drugs and Homicides



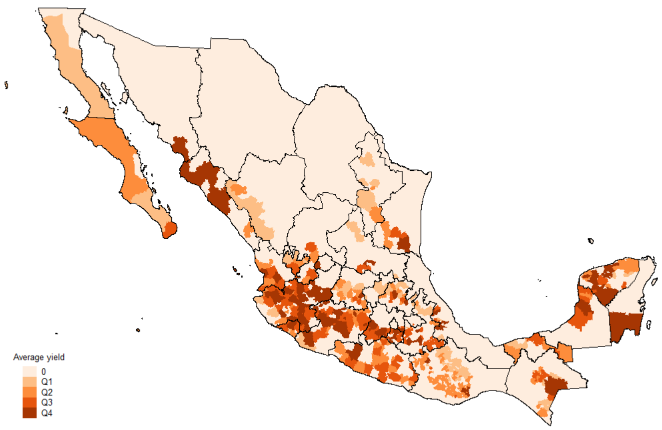
Note: Own calculation based on information from INEGI. Shaded areas represent the years of Calderon presidency. The Mexican War on Drugs started in 2007, the first year of his presidency. Bars summarizes the number of municipalities with at least one homicide registered in that year (right axis).

Figure 1.3: Cartel presence and expansion over time

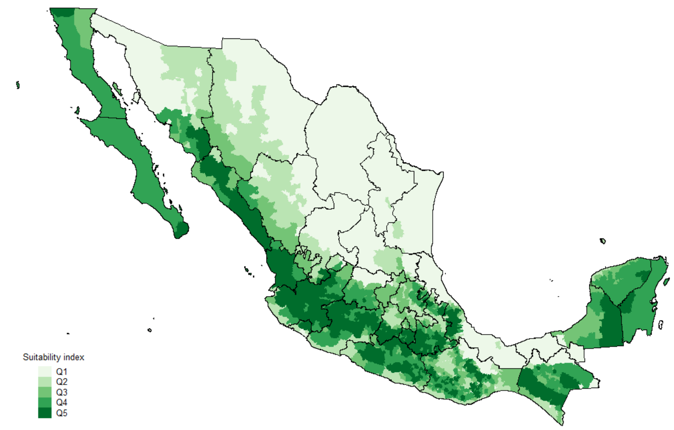


Note: Own calculation based on information from the Mapping Criminal Organizations project (Sobrino, 2020). The total number of municipalities with presence of cartels reported in the left vertical axis. The average number of cartels present in the municipalities reported in the right vertical axis. The dashed line represent the 90th percentile of the average presence of cartels.

Figure 1.4: Avocado presence, yield, and suitability index



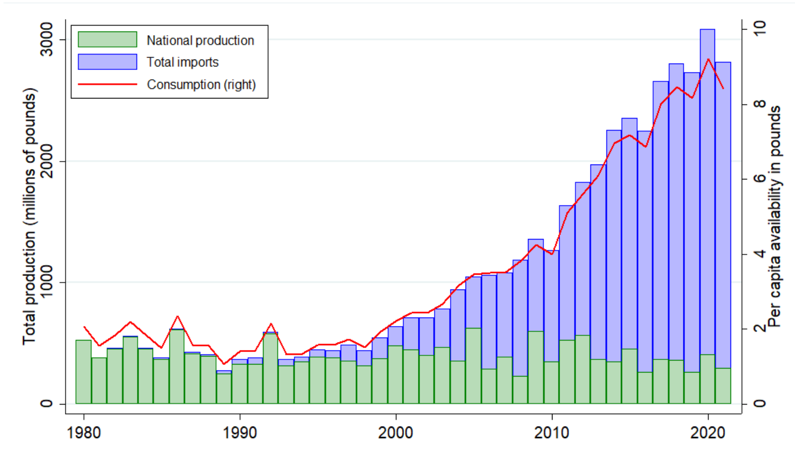
(a) Average yield



(b) Avocado suitability index

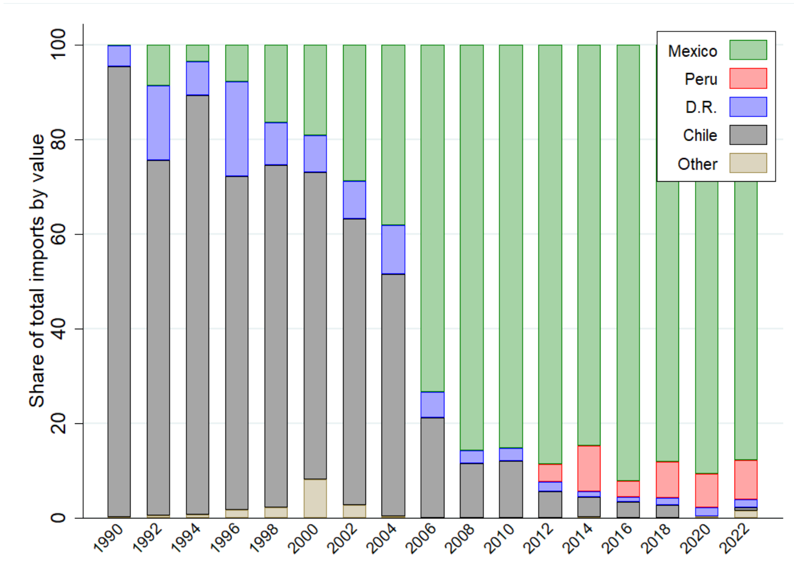
Note: Information from SIAP. Panel a presents the average yield of avocado in each one of the municipalities. Municipalities that do not produce avocado are label as “0” and producing municipalities are summarized by quantiles. Panel b presents the avocado suitability index as defined by equation 1.3 and it is summarized by quintiles. Darker shades of green indicates a higher suitability to grow avocado.

Figure 1.5: Production, imports, and availability of avocados in the United States



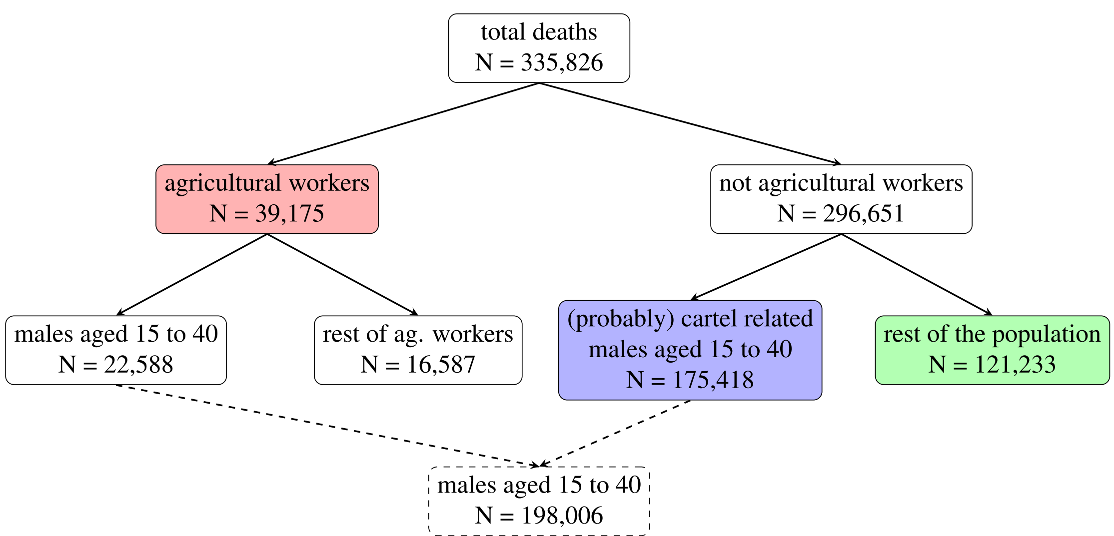
Note: Source: United States Department of Agriculture Economic Research Service (USDA-ERS). The green area represents the United States avocado production, and the blue area represents the imported avocado production. Red line represents the availability per capita in pounds of avocado as proxy for demand for avocado in the United States.

Figure 1.6: Avocado imports in the United States by country of origin



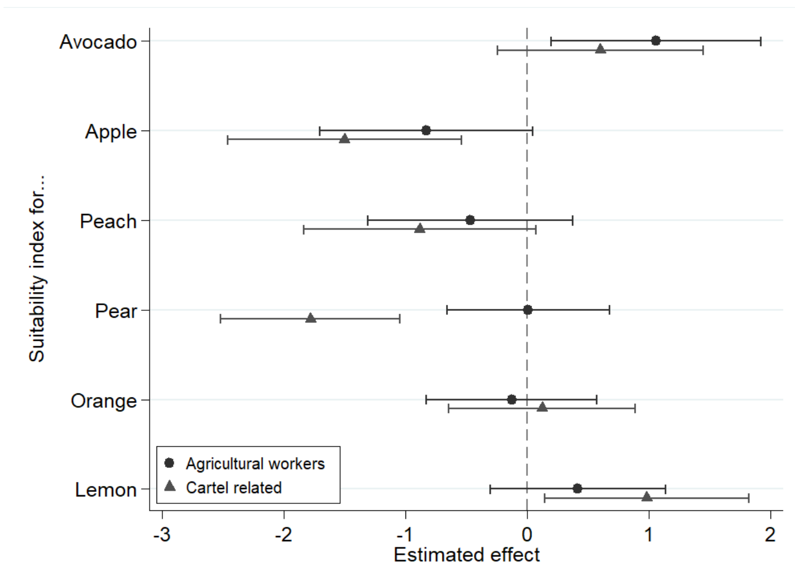
Note: Source: United States Department of Agriculture Foreign Agricultural Service's Global Agricultural Trade System (USDA-GAT). The share of imports is calculated for every year based on the value of total imports of each country.

Figure 1.7: Classification of homicides based on occupation, sex, and age



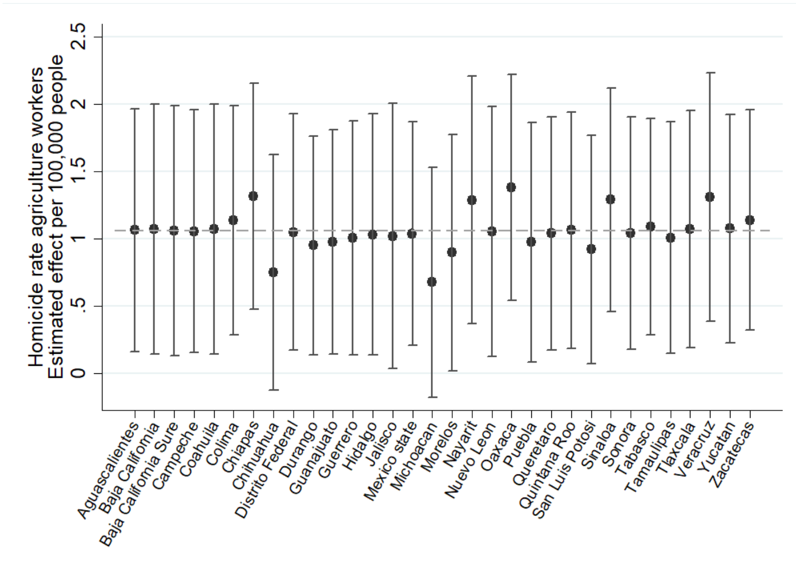
Note: The classification of homicides uses the information reported on the death certificates on occupation, age and gender. The first group consist of any agricultural worker, presented in red. The second group represent the probably cartel related deaths that consist of not agricultural workers, males between the ages of 15 to 40, presented in blue. The third group is the rest of the population, presented in green. The dashed group below represent all males between the ages of 15 to 40, regardless of their occupation.

Figure 1.8: Placebo test using suitability for growing other crops

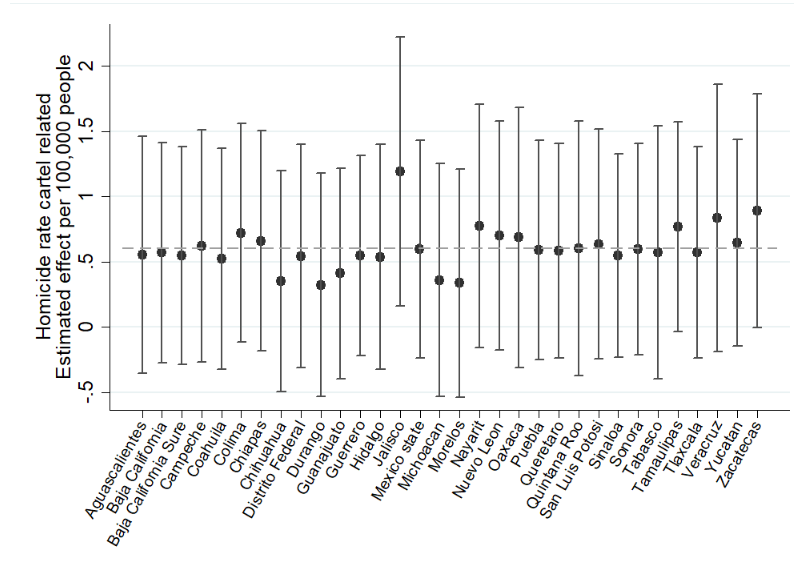


Note: Markers correspond to estimates of equation 1.1 where the explanatory variable is the interaction between the availability of avocados in the United States and a suitability index for growing the crop listed in the vertical axis. All estimates include municipality and time fixed effects, and control for the interaction of the municipal-level controls and year fixed effects. Block bootstrap standard error at the municipal level and 200 replications. Horizontal lines represent 95% confidence intervals .

Figure 1.9: Effects of potential to absorb avocado on homicide rate excluding each of the 32 states from the sample



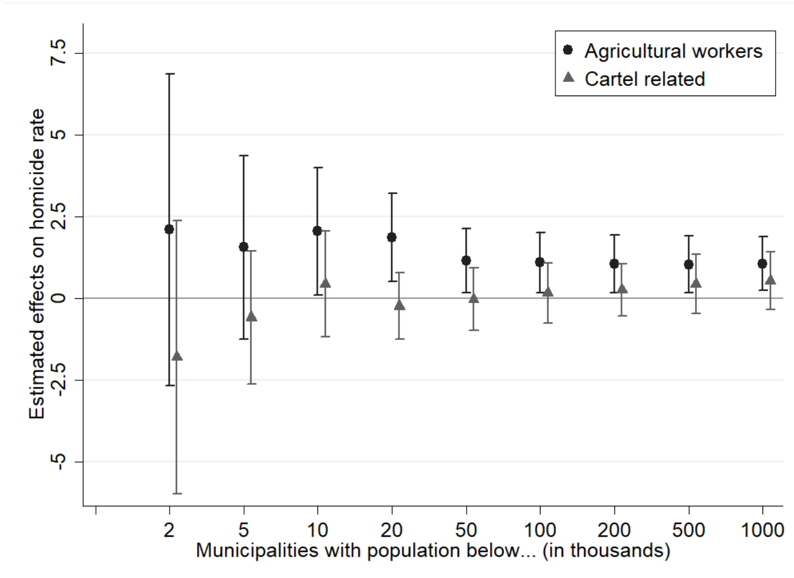
(a) Agricultural workers



(b) Cartel related

Note: Markers correspond to estimates coefficient of equation 1.1 excluding from the sample each one of the Mexican departments at a time, as reported on the horizontal axis. In Panel a presents the estimates on the homicide rate for agricultural workers and in Panel b presents the estimates on the homicide rate for potentially cartel related, i.e., non-agricultural male workers between the ages of 15 to 40. All the estimates include municipality and time fixed effects, and further control for the interaction of the municipal-level controls and year fixed effects. Horizontal dashed line reports the estimate using the full sample as reported in Table 1.4. Block bootstrap standard error at the municipal level and 200 replications. Vertical lines represent 95% confidence intervals.

Figure 1.10: Heterogeneous effects by population size



Note: Markers correspond to estimates of equation 1.1 using municipalities whose population in 2010 is below the value reported in the horizontal axis. All estimates include municipality and time fixed effects, and control for the interaction of the municipal-level controls and year fixed effects. Block bootstrap standard error at the municipal level and 200 replications. Vertical lines represent 95% confidence intervals .

Table 1.1: Descriptive statistics of the main variables and other panel variables

	Years	(1) Obs.	(2) Mean	(3) Std. Dev.
Panel A: Avocado suitability index and production				
Suitability x (log) US demand	2003 - 2019	41,752	0.478	0.37
Suitability x (log) US demand (standardized)	2003 - 2019	41,752	-0.000	1.000
Cultivated area (ha)	2003 - 2019	41,752	63.33	751.54
Quantity produced (thousand tons)	2003 - 2019	41,752	0.582	7.650
Value of production (million 2019 MXN)	2003 - 2019	41,752	10.122	144.05
Avocado incomde per capita (2019 MXN)	2003 - 2019	41,752	351.16	4,428.2
Panel B: Homicide data				
Homicide rate	2005 - 2019	36,794	16.96	34.0
Homicide rate agriculture workers	2005 - 2019	36,794	5.78	17.7
Homicide rate cartel related	2005 - 2019	36,794	6.00	15.2
Homicide rate rest of population	2005 - 2019	36,794	5.18	14.0
Panel C: Conflict data				
Homicide rate	2011 - 2019	22,104	26.25	41.5
Kidnapping rate	2011 - 2019	22,104	0.70	3.07
Extortion rate	2011 - 2019	22,104	2.26	6.19
Threat rate	2011 - 2019	22,104	28.17	58.7
Panel D: Cartel presence				
Presence of any DTO	2003 - 2019	41,684	0.212	0.4085
Presence of any traditional DTO	2003 - 2019	41,684	0.198	0.3987
Presence of any non-traditional DTO	2003 - 2019	41,684	0.096	0.2943
Total number of DTOs	2003 - 2019	41,684	0.710	2.13
Total number of traditional DTOs	2003 - 2019	41,684	0.484	1.25
Total number of non-traditional DTOs	2003 - 2019	41,684	0.227	1.08
Panel E: Labor market outcomes				
Unemployment rate	2005 - 2019	16,614	0.033	0.0260
Income (2019 x1000 MXN)	2005 - 2019	16,614	3.94	1.853
Income per hour (2019 MXN)	2005 - 2019	16,614	24.0	11.00
Income of agriculture (2019 x1000 MXN)	2005 - 2019	15,803	2.54	2.523
Income per hour of agriculture (2019 MXN)	2005 - 2019	15,803	15.7	21.08
Panel F: Economic composition				
Employment in agriculture	2005 - 2019	16,614	0.278	0.2563
Employment in construction	2005 - 2019	16,614	0.081	0.0630
Employment in indsutry	2005 - 2019	16,614	0.146	0.1157
Employment in services	2005 - 2019	16,614	0.316	0.1695
Employment in commerce	2005 - 2019	16,614	0.165	0.0847
Panel G: Fiscal capacity				
Revenues from taxes	2003 - 2019	33,916	0.037	0.0549
Taxes from production taxes	2003 - 2019	13,737	0.070	0.1694
Revenue from production taxes	2003 - 2019	13,737	0.002	0.0066
Expenses from public investment	2003 - 2019	35,337	0.349	0.1771
Expenses from personal expenses	2003 - 2019	35,902	0.267	0.1230

Note: Crime rates reported per 100,000 people. Availability of the data is reported in the column “Years.”

Table 1.2: Descriptive statistics of the cross-sectional variables

	(1) Obs.	(2) Mean	(3) Std. Dev.
Panel I: Geographic characteristics			
Suitability index	2,456	0.297	0.216
Altitude	2,456	1,301	821.7
Water availability	2,456	1,049	581.4
Soil index	2,456	3.94	2.55
Panel II: Control variables			
Population in 2000	2,443	39,903	119,428
Social gap index 2000	2,442	0.000	1.000
Gini coefficient 2000	2,453	0.461	0.069
Distance to municipal main area (in degrees)	2,452	1.034	0.718
Presence of cartels 1990-2001	2,452	0.094	0.292

Note: Panel I presents descriptive statistics of the avocado suitability index and the variables used to construct it. Panel II presents the control variables.

Table 1.3: Avocado suitability index

VARIABLES	(1) Avocado (logit)	(2) Avocado (OLS)	(3) Yield	(4) I.H.S. Yield
Water availability	0.003***	0.000***	0.003***	0.001***
	(0.001)	(0.000)	(0.000)	(0.000)
Water availability squared	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Soil index	0.256***	0.042***	0.453***	0.133***
	(0.064)	(0.011)	(0.083)	(0.028)
Soil index squared	-0.027***	-0.004***	-0.041***	-0.012***
	(0.007)	(0.001)	(0.009)	(0.003)
Altitude	0.004***	0.001***	0.004***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Altitude squared	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	2,392	2,456	2,456	2,456
R-squared		0.171	0.187	0.182
Physiographic regions controls	✓	✓	✓	✓
Pseudo or adjusted R-squared	0.163	0.163	0.180	0.175

Note: Columns report the estimates from equation 1.2. Outcome variable is reported in the column header. Inverse hyperbolic sine (I.H.S.) transformation used instead of the logarithm due to the presence of zeros. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.4: Effects of potential to absorb demand for avocado on homicides

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Homicide rate								Avocado value	
	Homicides	Homicides	Agricultural worker	Agricultural worker	Cartel related	Cartel related	Rest of population	Rest of population	per capita (1st-stage)	per capita (1st-stage)
	Panel A: reduce form									
Suitability × US demand	1.366	1.960	1.612	1.060	0.052	0.603	-0.297	0.297	1.326	1.397
	(0.929)	(0.918)**	(0.428)***	(0.452)**	(0.462)	(0.441)	(0.372)	(0.311)	(0.236)***	(0.258)***
	[1.578]	[1.124]*	[0.640]**	[0.550]*	[0.728]	[0.490]	[0.463]	[0.357]	[0.275]***	[0.278]***
	Panel B: 2SLS IV regression (second-stage)									
Avocado value per capita	1.030	1.404	1.215	0.759	0.039	0.432	-0.224	0.213		
	(0.732)	(0.712)**	(0.389)***	(0.336)**	(0.343)	(0.321)	(0.275)	(0.251)		
	[1.116]	[0.813]*	[0.471]***	[0.364]**	[0.545]	[0.358]	[0.365]	[0.267]		
Observations	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529
Number of municipalities	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438
Mean dep. var.	16.96	16.96	5.78	5.78	6.00	6.00	5.18	5.18	0.35	0.35
Kleibergen-Paap F-stat.									29.09	29.65
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Panel A presents the reduce form estimates. Panel B present the second stage estimates of a 2SLS. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The The column header reports the dependent variable and the bottom panel reports its mean in the sample. All the outcomes are rates per 100,000 people. Cartel related refers to non-agricultural male workers between the ages of 15 and 40. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. Conley (1999) standard errors clustered using a 500 Km radius are reported in square brackets. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.5: Potential to absorb demand for avocado and actual production of avocado

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	in levels								in logs (IHS)			
	Area cultivated	Area cultivated	Volume produced	Volume produced	Value produced	Value produced	Value per capita	Value per capita	Area cultivated	Volumme produced	Value produced	Value per capita
Suitability × log(US demand)	130.45***	132.17***	1.27***	1.29***	38.47***	39.28***	1,257.4***	1,319.5***	0.671***	0.198***	0.594***	0.955***
	(23.85)	(24.83)	(0.25)	(0.26)	(8.85)	(9.09)	(235.5)	(258.0)	(0.061)	(0.020)	(0.043)	(0.117)
Observations	41,752	41,446	41,752	41,446	41,752	41,446	41,752	41,446	41,446	41,446	41,446	7,701
Number of municipalities	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438	2,438	2,438	2,438	688
Mean Dep. Var.	63.33	63.33	0.58	0.58	8.13	8.13	351.2	351.2	0.759	0.093	0.260	4.363
F-stat	26.11	27.20	22.86	24.26	18.02	18.51	26.49	26.97	125.29	96.53	195.54	62.05
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓	✓	✓	✓	✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Column (9), (11), and even columns further control for the interaction of the municipal-level controls and year fixed effects. Column header reports the dependent variable. In Columns (1) to (8), the dependent variables are in levels, while in Columns (9) to (12), they are in logs using the Inverse Hyperbolic Sine (I.H.S.) transformation due to the presence of zeros. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.6: Robustness to time-varying unobserved heterogeneity using rolling estimates

	(1) FE	(2) FD	(3) TFD	(4) RFD (cons)	(5) RFD (nocons)	(6) RTFD (cons)	(7) RTFD (nocons)
Panel A: Homicide rate							
Suitability × US demand	1.960**	7.361***	13.822***	8.845*	7.948	14.728***	14.544**
	(0.932)	(2.377)	(4.685)	(4.734)	(4.860)	(4.757)	(4.913)
Panel B: Homicide rate for agricultural workers							
Suitability × US demand	1.060**	2.732**	5.073*	4.195	3.807	6.225**	5.870*
	(0.439)	(1.300)	(2.631)	(2.539)	(2.453)	(2.849)	(2.811)
Panel C: Homicide rate for potentially cartel related							
Suitability × US demand	0.603	1.187	1.767	0.551	0.053	0.168	0.200
	(0.433)	(1.209)	(2.389)	(1.627)	(1.781)	(1.652)	(1.909)
Panel D: Homicide rate for the rest of the population							
Suitability × US demand	0.297	3.442***	6.982***	3.279	3.164	7.845***	7.903***
	(0.345)	(1.182)	(2.379)	(2.119)	(2.053)	(2.168)	(2.078)

Note: Each column reports estimates of equation 1.1. Column headers report the estimator used: Fixed Effects (FE), First Difference (FD), Twice-First Difference (TFD), the Rolling First Difference (RFD), and the Rolling Twice-First Difference (RTFD) (Millimet and Bellemare, 2023). Columns (5) and (7) rolling estimates include a constant term. Columns (6) and (8) rolling estimates do not include a constant term. Potential to absorb demand for avocado variable is standardized in all specifications. All columns include municipality and year fixed effects. Dependent variable in Panel A is homicide rate in the municipality; in Panel B is the homicide rate of agricultural workers; in Panel C is the homicide rate for individuals potentially cartel related; and in Panel D the homicide rate for the rest of the population. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.7: Effects of potential to absorb demand for avocado on homicides of males between the ages of 15 to 40

	(1) Agricultural worker	(2) Agricultural worker	(3) Ag. workers males 15-40	(4) Ag. workers males 15-40	(5) Ag. workers rest	(6) Ag. workers rest	(7) males 15-40	(8) males 15-40	(9) Cartel related	(10) Cartel related
Suitability $\times$ log(US demand)	1.612***	1.060**	1.225***	0.735**	0.387	0.325	1.276**	1.339**	0.052	0.603
	(0.428)	(0.452)	(0.280)	(0.310)	(0.239)	(0.229)	(0.611)	(0.621)	(0.462)	(0.441)
Observations	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529
Number of municipalities	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438
Mean Dep. Var.	6.047	6.047	3.252	3.252	2.796	2.796	10.099	10.099	6.847	6.847
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls $\times$ Year Fe		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The The column header reports the dependent variable and the bottom panel reports its mean in the sample. All the outcomes are rates per 100,000 people. Male individuals between the ages of 15 to 40 includes both agricultural and non-agricultural workers. Cartel related refers to non-agricultural male workers between the ages of 15 and 40. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.8: Effects of potential to absorb demand for avocado on homicides of rural residents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Rural Homicides	Rural Homicides	Ag. worker Rural	Ag. worker Rural	Cartel related Rural	Cartel related Rural	Rest of population Rural	Rest of population Rural
Suitability × log(US demand)	0.611	0.704	1.068***	0.712**	-0.291	0.005	-0.166	-0.013
	(0.598)	(0.562)	(0.354)	(0.357)	(0.303)	(0.242)	(0.272)	(0.231)
Observations	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529
Number of municipalities	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438
Mean Dep. Var.	9.116	9.116	4.301	4.301	2.422	2.422	2.393	2.393
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. All the outcomes are rates per 100,000 people. Cartel related refers to non-agricultural male workers between the ages of 15 and 40. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.9: Effects of potential to absorb demand for avocado on criminal activities

	(1) Homicides	(2) Homicides	(3) Threats	(4) Threats	(5) Extortions	(6) Extortions	(7) Kidnappings	(8) Kidnappings
Suitability × log(US demand)	5.791*	5.159*	-1.192	0.448	-1.644***	-1.333**	-0.214	-0.251
	(3.157)	(2.934)	(4.130)	(3.670)	(0.548)	(0.534)	(0.238)	(0.218)
Observations	22,104	21,942	22,104	21,942	22,104	21,942	22,104	21,942
Number of municipalities	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438
Mean Dep. Var.	26.250	26.250	28.175	28.175	2.260	2.260	0.701	0.701
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. All the outcomes are rates per 100,000 people. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.10: Effects of potential to absorb demand for avocado on cartel presence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Presence of cartels						Number of cartles					
	any	any	traditional	traditional	non-traditional	non-traditional	any	any	traditional	traditional	non-traditional	non-traditional
Suitability × log(US dmand)	0.067***	0.085***	0.061***	0.082***	0.043***	0.055***	0.307***	0.398***	0.186***	0.242***	0.122**	0.155***
	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)	(0.011)	(0.092)	(0.088)	(0.044)	(0.038)	(0.058)	(0.059)
Observations	41,684	41,378	41,684	41,378	41,684	41,378	41,684	41,378	41,684	41,378	41,684	41,378
Number of municipalities	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434
Mean Dep. Var.	0.212	0.212	0.198	0.198	0.096	0.096	0.710	0.710	0.484	0.484	0.227	0.227
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. In Columns (1) to (3), the dependent variable is a dummy variable that indicates presence of any cartel, traditional, and non-traditional cartels only, respectively. In Columns (4) to (6), the dependent variable is the number of cartels present in the municipality, of traditional, and of non-traditional. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.11: Heterogeneous effects on homicides by “contemporaneous” cartel presence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Dep. var: Homicide rate agricultural workers						Dep. var: Homicide rate cartel related					
	any	any	traditional	traditional	non-traditional	non-traditional	any	any	traditional	traditional	non-traditional	non-traditional
Suitability × US demand	1.803***	1.201**	1.865***	1.259**	1.470***	0.943*	-0.161	0.867*	-0.048	0.941*	-0.802*	0.039
	(0.499)	(0.558)	(0.490)	(0.549)	(0.441)	(0.497)	(0.507)	(0.488)	(0.504)	(0.483)	(0.453)	(0.447)
Suitability × US demand × cartel (t-1)	-0.213	-0.148	-0.303	-0.224	0.268	0.221	0.227	-0.285	0.109	-0.388*	1.612***	1.044***
	(0.210)	(0.196)	(0.203)	(0.189)	(0.231)	(0.204)	(0.237)	(0.228)	(0.251)	(0.228)	(0.355)	(0.306)
Observations	36,734	36,469	36,734	36,469	36,734	36,469	36,734	36,469	36,734	36,469	36,734	36,469
Number of municipalities	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434
Mean Dep. Var.	5.779	5.779	5.779	5.779	5.779	5.779	5.999	5.999	5.999	5.999	5.999	5.999
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable, the type of cartel, and the bottom panel reports its mean in the sample. In Columns (1) to (6) the dependent variable is the homicide rate of agricultural workers, and in Columns (7) to (12) is the homicide rate of potentially cartel related, i.e., non-agricultural male workers between the ages of 15 and 40. All the outcomes are rates per 100,000 people. $cartel_{t-1}$ is a dummy variable that takes value of 1 if there was a cartel present in the previous years. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.12: Heterogeneous effects on homicides by cartel presence in the 1990s

	(1) Homicides	(2) Homicides	(3) Agricultural worker	(4) Agricultural worker	(5) Cartel related	(6) Cartel related	(7) Rest of population	(8) Rest of population
Suitability × US demand	0.647	2.132**	1.907***	1.477***	-0.655	0.354	-0.606	0.301
	(0.951)	(1.019)	(0.428)	(0.502)	(0.450)	(0.440)	(0.375)	(0.401)
Suitability × US demand × cartel 90s	5.174**	-0.998	-2.086***	-2.621***	4.901***	1.545	2.359***	0.077
	(2.052)	(1.986)	(0.767)	(0.888)	(1.104)	(1.056)	(0.570)	(0.540)
Observations	36,734	36,469	36,734	36,469	36,734	36,469	36,734	36,469
Number of municipalities	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434
Mean Dep. Var.	16.956	16.956	5.779	5.779	5.999	5.999	5.177	5.177
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. All the outcomes are rates per 100,000 people. $cartel\ 90s$ is a dummy variable that takes value of 1 if there was a cartel present at least one year between 1990 and 2000. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.13: Heterogeneous effects on homicides by cartel presence of multiple cartels in the previous period

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Dep. var: Homicide rate agricultural workers						Dep. var: Homicide rate cartel related					
	any	any	traditional	traditional	non-traditional	non-traditional	any	any	traditional	traditional	non-traditional	non-traditional
Suitability × US demand	1.604***	1.000*	1.774***	1.159**	1.427***	0.916*	-0.549	0.509	-0.376	0.626	-0.917**	-0.091
	(0.495)	(0.519)	(0.483)	(0.500)	(0.458)	(0.477)	(0.507)	(0.476)	(0.503)	(0.467)	(0.458)	(0.437)
Suitability × US demand × one cartel_$t-1$	0.120	0.158	-0.170	-0.091	0.356	0.316	0.418*	0.003	0.284	-0.109	1.013***	0.596**
	(0.269)	(0.279)	(0.243)	(0.249)	(0.295)	(0.290)	(0.233)	(0.219)	(0.225)	(0.207)	(0.295)	(0.295)
Suitability × US demand × two+ cartel_$t-1$	-0.034	0.028	-0.202	-0.120	0.438	0.284	0.761***	0.138	0.611**	0.017	3.216***	2.412***
	(0.239)	(0.287)	(0.238)	(0.296)	(0.353)	(0.375)	(0.278)	(0.276)	(0.295)	(0.281)	(0.582)	(0.596)
Observations	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529
Number of municipalities												
Mean Dep. Var.	5.779	5.779	5.779	5.779	5.779	5.779	5.999	5.999	5.999	5.999	5.999	5.999
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable, the type of cartel, and the bottom panel reports its mean in the sample. In Columns (1) to (6) the dependent variable is the homicide rate of agricultural workers, and in Columns (7) to (12) is the homicide rate of potentially cartel related, i.e., non-agricultural male workers between the ages of 15 and 40. All the outcomes are rates per 100,000 people. $onecartel_{t-1}$ is a dummy variable that takes value of 1 if there was only one cartel present in the previous year, and $two + cartel_{t-1}$ is a dummy that takes the value of 1 if there were two or more cartels present in the previous year. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.14: Effects of potential to absorb demand for avocado on labor market outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
							Ag. workers			
	Unemployment	Unemployment	(log) income	(log) income	income x hr	income x hr	(log) income	(log) income	income x hr	income x hr
Suitability × log(US demand)	0.003***	0.004***	0.023	0.026	0.824	0.745	0.035	0.018	1.806**	1.506
	(0.001)	(0.001)	(0.024)	(0.022)	(0.582)	(0.540)	(0.033)	(0.032)	(0.858)	(1.014)
Observations	16,614	16,541	16,611	16,538	16,614	16,541	15,449	15,379	15,803	15,730
Number of municipalities	1,793	1,779	1,793	1,779	1,793	1,779	1,783	1,769	1,788	1,774
Mean Dep. Var.	0.033	0.033	1.232	1.234	23.956	23.984	0.585	0.588	15.699	15.734
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.15: Effects of potential to absorb demand for avocado on labor composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Primary		Secondary				Tertiary			
	Agriculture	Agriculture	Construction	Construction	Manufacturing	Manufacturing	Services	Services	Commerce	Commerce
Suitability × log(US demand)	0.018	0.017	0.001	-0.000	-0.017***	-0.017***	-0.001	0.000	-0.001	-0.000
	(0.012)	(0.012)	(0.003)	(0.003)	(0.004)	(0.005)	(0.008)	(0.009)	(0.004)	(0.004)
Observations	16,614	16,541	16,614	16,541	16,614	16,541	16,614	16,541	16,614	16,541
Number of municipalities	1,793	1,779	1,793	1,779	1,793	1,779	1,793	1,779	1,793	1,779
Mean Dep. Var.	0.278	0.278	0.081	0.081	0.146	0.147	0.316	0.316	0.165	0.165
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 1.16: Effects of potential to absorb demand for avocado on fiscal capacity

	(1) Revenues from taxes	(2) Revenues from taxes	(3) Revenues from prod. taxes	(4) Revenues from prod. taxes	(5) Taxes from prod. taxes	(6) Taxes from prod. taxes	(7) Expenses from public inv.	(8) Expenses from public inv.	(9) Expenses from personal exp.	(10) Expenses from personal exp.
Suitability × log(US demand)	0.003***	0.002*	0.001*	0.001*	0.005	0.004	-0.005	-0.005	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.006)	(0.005)	(0.005)	(0.003)	(0.004)
Observations	33,916	33,704	13,737	13,649	13,737	13,649	35,337	35,092	35,902	35,649
Number of municipalities	2,419	2,401	1,922	1,908	1,922	1,908	2,440	2,422	2,440	2,422
Mean Dep. Var.	0.037	0.037	0.002	0.002	0.070	0.070	0.349	0.349	0.267	0.267
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. Block bootstrap standard error at the municipal level and 200 replications are reported in parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Chapter 2: Books and Bushes: Schooling Decisions and Coca Production in Colombia

2.1 Introduction

This paper explores the relation between agricultural shocks and schooling decisions. There are many papers that relate booms in legal markets with schooling decisions. Examples abound such as the coffee boom in Brazil, where poorer children were withdrawn from school (see for instance Kruger, 2007) Yet, there are less papers that relate booms in illegal markets with schooling decisions. In particular, there are just a handful of papers that related the presence of coca leaves with schooling decisions (Angrist and Kugler, 2008; Dammert, 2008; Sviatschi, 2022).

In this paper I look at the effect of coca crops presence on school participation. I focus on the case of Colombia where coca fields were increasing at an average annual rate of 25 percent, from 2013 to 2016. I take advantage of the temporary nature of the fumigation campaigns of herbicide as a strategy to eradicate coca crops to assess whether school-age individuals decide to remain in school. To do so, I first explore whether there is variation on school outcomes; in particular, I look at the effect of coca presence on enrollment and dropout rate. Then, I focus the attention on what happen to school-age individuals that decide to remain and to leave school. I look at school attainment measures to document the changes in the patterns of individuals that

[0]This paper received support from the Institute for Humane Studies, grant no. IHS016606.

decide to remain in school. Finally, I look at (legal) agricultural activities to shed light on whether the presence of coca crops is related to changes in the agricultural sector.

The ideal experiment would compare school outcomes in municipalities that are identical in all respects except exposure to coca fields. To approximate such an experimental ideal, I use municipal variation on the share of area cultivated with coca fields. I take advantage of the spike in the presence of coca fields starting 2013, which can be characterized as a boom in coca leaf production and whose production is labor intensive. Harvesting coca crops demands time and physical effort, but it does not require different skills than the ones needed to harvest legal crops.[1] Either children or adults can work on the harvesting activities of coca where a plot is harvested a minimum of 4 times a year. Previous studies have related coca production with children and teenage work, both in Colombia (Angrist and Kugler, 2008; Rodríguez, 2020) and in Peru (Sviatschi, 2022).

As part of the 'war on drugs', the United States and the Colombian government launched *Plan Colombia* in late 1999, a policy-package aimed at reducing the country cocaine supply. The strategies implemented included the forced eradication of coca crops through the aerial spraying campaigns of herbicide. However, after the World Health Organization found that the active component of the herbicide sprayed (namely glyphosate) is "probably carcinogenic to humans," the Colombian government under the then-President Santos suspended the aerial spraying of coca fields in late 2015.

Aerial spraying of herbicide is a short-term strategy as it does not have a permanent effect. Areas fumigated can recover in less than one year. Thus, they could be fumigated multiple

[1] Note, coca leaf is the primary input in the production of cocaine. Turning coca leaves into coca paste and then into cocaine is less labor intensive and require from additional inputs, mostly chemicals.

times over the years or within the same year. Fumigation only occurred in areas affected with coca presence yet not all the areas affected are indeed sprayed. Therefore, I look at exogenous variation in the policy using both cross-sectional and longitudinal variation. I calculate the share of area affected with coca presence in a given year, and I instrument it using a dummy variable for whether the area was fumigated or not in that year. Notice, the last year I can study is 2016 as the fumigation policy was suspended in 2015.

Intuitively, my instrument picks two consecutive periods and compare an area that was not sprayed with an area that was sprayed. Using the policy suspension, I exploit the temporal variation as in a difference-in-differences framework. On the other hand, I exploit the cross-sectional variation as not all the areas with coca presence are sprayed which lead to a fuzzy design. Notice, spraying missions are centrally planned and administered by the Anti-Narcotic Police after aerial recognition of cultivated areas, then I can hardly claim that fumigation is random. Nonetheless, I present some evidence suggesting that the missions are determined by geographic characteristics rather than by coca presence. Therefore, I claim that the spraying mission are decided quasi-randomly.

Both the temporal and cross-sectional variation lead to a difference-in-differences framework, however my setting differ from that of the 'classical' difference-in-differences so I borrow from the literature the Fuzzy Difference-in-Differences methodology (de Chaisemartin and D'Haultfœuille, 2018). First, there are areas where the government stop spraying before the suspension of the fumigation campaigns, which could be interpreted as staggered adoption of the 'treatment.' Then, the aerial spraying missions did not have a permanent effect as municipalities do not remain exposed to the treatment at all times afterwards. These challenges are tackled by the fuzzy difference-in-differences design by comparing groups that adopt and remove the treatment while controlling for the timing of the adoption/removal of the policy, i.e., whether or not the municipality is being sprayed.

My empirical results are consistent with the expected values of a negative effect of coca presence on school outcomes. My findings suggest that the increasing presence of coca crops is significantly correlated with more students dropping out from high school, while I do not find any effect on the other grades nor in the enrollment rate. I find that on average dropout rate increases by 0.33 standard deviations in areas with presence of coca crops, which represent a decrease of almost 30% relative to the 3.15% mean high school dropout rate. In addition, the retention rate decreases, although its magnitude is much smaller (0.1 percentage points).

I look at the performance of the students that decided to stay in school using data on the rate of students that either pass or fail and repeat a level in primary, middle, and high school. My results suggests that the increasing presence of coca fields have almost no effect on the students that remained in school, except for those in high school, where there is a negative effect on the rate of students that approved of their level. This means that students that did not drop out from school are not fulfilling the requirements of their grades. Moreover, I use information from the standardized tests—Saber5, Saber9, and Saber11 that students take at the end of primary, middle, and high school respectively—to assess the effect of coca presence on learning outcomes. I do not find any effect on the tests scores except for the mathematics component in primary school which present a positive effect. Even though this results is consistent with previous studies that show that students that remain in school tend to perform better, it is unexpected as there were no changes in dropout nor enrollment in that level.

Finally, to shed light on an outside option for the school-age students, I explore changes in the production of other legal crops. I use data on the area cultivated and harvested in each municipality, as well as its ratio. These measures allow me to evaluate whether coca crops crowd out legal crops. My results suggests that the coca presence has no effect either on the area sown nor in the area harvested, which could be interpreted as the agricultural labor market being in equilibrium. Since the only change is in the area cultivated with illegal crops and not on the area with licit crops, I interpret these results as suggestive evidence that, on the one hand, there is not a crowd out effect of the coca crops on the legal crops, and on the other hand, that school-age individuals that decide to drop out are working on coca related activities. Notice, I cannot claim that this is in fact the case, as I do not have data to support it, but I cannot rule it out either.

A limitation of my analysis is the level of aggregation of the school outcome variables. They are at the year-municipal level while the decision of enrolling or dropping out of school is at the individual level. Moreover, tests are taken by individual students and this level of aggregation could mix top-performing schools with bottom-performing in the same municipality. Similarly, both the coca presence and the spraying data are at the year-municipal level, which do not allow me to capture short-term dynamics nor find a more precise effect. For instance, the presence of coca crops could increase at the end of the year and the student might dropout in the middle of it because families decide to migrate from the rural area to the urban area or to a different municipality. On the other hand, I cannot distinguish the effect by grades, or within school levels. For instance, the positive enrollment in one grade is averaged with the negative enrollment in another grade within the same school level.

This study connects with the literature related to eradication of coca crops through aerial spraying. Some of these studies have found evidence of small but positive effects of the aerial spraying campaigns aimed at reducing coca cultivation (Abadie et al., 2015; Cote, 2019; Ibanez and Carlsson, 2010; Mejía, Restrepo and Rozo, 2017; Prem, Vargas and Mejía, 2023; Reyes, 2014), although there is one study that suggests that an increase in forced eradication leads to an increase in the area cultivated (Reyes, 2014). Other studies have found that the presence of coca fields is related to children and teenager work, both in Colombia (Angrist and Kugler, 2008; Rodríguez, 2020) and in Peru (Sviatschi, 2022). While others have also related aerial spraying with more displacement (Espinosa, 2010), and with health outcomes, such as of medical consultations related to dermatological and respiratory illnesses, as well as the number of miscarriages (Camacho and Mejía, 2017)

This paper also contributes to the studies that explore the relation between exposure to illegal markets and education and labor decisions (Dammert, 2008; Dávalos and Dávalos, 2020; Ibáñez, 2010). There is one paper that exploits the boom in Colombian illegal gold mining, finding that school attendance decreases as well as the probability of child work increases (Santos, 2018). There are two papers that uses the presence of coca fields in Peru and relate it with child work and schooling decisions; one uses the shift of production from Peru to Colombia (Angrist and Kugler, 2008) while the other exploit the exposure to illegal labor markets (Sviatschi, 2022).

The studies closer to mine are those of Rodríguez (2020) and of Prem, Vargas and Mejía (2023). The former investigates the effect of the aerial spraying policy on education decisions using survey data in areas affected with coca presence; although, their study cover the period before the boom in coca fields. On the other hand, the latter study exploit the unintended consequences of a policy announcement where the coca growing farmers would receive incentives for voluntarily crop substitution. Their results suggests that the surge in coca fields is differentially higher in areas with presence of illegal armed groups.

To the best of my knowledge, this is the first paper that studies the consequences of the coca boom of the beginning of the 2010s. In addition, this is the first paper that studies the removal of the aerial spraying strategy and that uses the standardized test scores Saber 5 and Saber 9 to show heterogeneous effects of learning outcomes across grades while relating them to coca presence.

The rest of the paper is organized as follows. Section 2.2 provides some background on the Colombian case. Section 2.3 summarize the data sources and describe the research design. Section 2.4 describes the empirical strategy to estimate the effect of coca presence on school outcomes. Section 2.5 reports the main findings on school outcomes, school efficiency, and school attainment. Finally, Section 2.6 concludes.

2.2 Institutional Context and Background

2.2.1 Dynamics of coca presence in Colombia

Coca leaf is the main ingredient of cocaine, a powerful addictive drug, whose mayor producer is Colombia with 70 percent of the world supply[2] and approximately 89 percent of the cocaine seized and subjected to laboratory analysis in the United States in 2019.[3] This is a permanent or multi-annual crop that can be harvested a minimum of 4 times a year (4.6 times on average), and the cost of producing it decreases over time. Harvesting coca leaves is a labor-intensive process and demands time and physical effort, a similar set of skills than the ones needed to harvest licit crops. Workers are needed for plowing the earth, plantation, and fumigation besides of harvesting, where there is more demand (Dammert, 2008; Rodríguez, 2020). One hectare[4] (ha) of coca bushes produces on average 1,217 kilograms of fresh coca leaves per

[2] See https://www.semana.com/mundo/articulo/produccion-de-coca-crece-en-el-mundo-y-colombia-aporta-el-70/621010.

[3] See https://www.state.gov/bureau-of-international-narcotics-and-law-enforcement-affairs-work-by-country/colombia-summary/.

[4] One hectare (ha) is equivalent to approximately 2.5 acres.

harvest, and 1 kilogram of coca leaves can produce between 1.1 and 1.4 grams of cocaine. Depending on multiple factors, such as density of sowing, soil nutrients, and weather condition, it can present different yield , but it is estimated between 5 and 6 kilograms of cocaine per hectare per year (DIRAN, 2014; Mejía and Posada, 2008).

The revenue from this stage of production depends on how many kilograms of coca leaves one person can harvest and carry. In 2016, the price for one kilogram of coca leaves was on average \$2,900 Colombian pesos (roughly[5] \$0.95 US dollars), although the price varies between regions. It is estimated that a person could receive about \$960 US dollar a year (UNODC, 2017). For a farmer harvesting its own plot, the National Police has estimated an income between 2 and 3 monthly minimum wages which is slightly above the region's median income (DIRAN, 2014). Notice, farmers decide to grow coca as there is no possibility of making a living from legal forms of agriculture (Ibáñez, 2010), either because they are located far from city centers (Zuleta, 2019) or because they have less access to formal credits for their licit crops (Dávalos and Dávalos, 2020), among other reasons.

Since the late 1990s, Colombia has aggressively pursued coca eradication to reduce the land devoted to it (Reyes, 2014). In 1999, the United States and the Colombian governments launched *Plan Colombia*, a \$7.5 billion co-financed policy-package, aimed at reducing 50 percent the cultivation over a period of six years. Aerial spraying of herbicide (fumigation) over coca fields was the main instrument to eradicate the crop, along with manual eradication either forced or voluntary (Abadie et al., 2015; Zuleta, 2019). Notice, Colombia is the only country that allowed aerial spraying for counter-drug purposes (Isacson, 2019).

Spraying missions are centrally planned by the Anti-Narcotic Police (DIRAN) after aerial

[5] \$1 US dollar is equivalent to \$3,052 Colombian pesos using 2016 exchange rate.

reconnaissance of cultivation areas. This strategy allows operating in remote and insecure areas where manual eradication is cost prohibitive or too dangerous (Abadie et al., 2015). Spraying missions are made through small aircraft spraying glyphosate, the active ingredient in the herbicide RoundUp (Isacson, 2019), using modified military planes (fumigation equipment instead of weapons). Spraying aircraft are protected and escorted by the National Police with armed helicopters to protect fumigation planes from armed attacks from the ground. These helicopters have a range of 80 miles from the airport, beyond that range missions are not canceled but fumigation planes must go unprotected (Reyes, 2014).

According to DIRAN, spraying glyphosate is 90 percent effective.[6] To estimate the area affected with coca fields presence, the monitoring system SIMCI uses an 8.8 percent survival rate of a sprayed crop (UNODC, 2017). Once a field is sprayed, the land takes six to eight months to regenerate to a point at which coca can be grown again. However, if it rains or if growers wash the crops immediately after the spraying, the effect of the spraying is mitigated and the land recovers much faster. If done right, the herbicide can be "washed out" from plants, as long as this is done right after the spraying has taken place (Prem, Vargas and Mejía, 2023; UNODC, 2014). Given this, plots are likely to be sprayed more than once in a given year. Spraying missions were suspended in October 2015 by the Colombian government after the World Health Organization classified glyphosate as "probably carcinogenic to humans" in March of the same year.[7]

Figure 2.1 presents the dynamics of the coca crops in Colombia. It shows the total area under coca bush cultivation and the total area eradicated. Notice, the latter is higher in magnitude than the former for various years during the 2010s. This is evidence of areas being sprayed

[6] Personal communication with official from DIRAN, July 2019

[7] After the suspension, Colombia's Constitutional Court, one of the three highest judicial review authorities, placed restrictions on any use of glyphosate from aircraft, but did not ban the practice (Isacson, 2019).

more than once in a given year. After the suspension of the missions in 2015, the eradication is done through 'manual' eradication by pulling coca bushes from the ground. The area with coca presence increased from 48,000 ha in 2013 to 146,000 ha in 2016. Previous studies have suggested that the peace dialogue between the Colombian government and the FARC guerrilla play a role in this increment as there is a non-trivial relation between the latter group and the illicit crops (López et al., 2019; Zuleta, 2019). One study suggest an anticipatory effect generated by the announcement of material incentives for voluntary crop substitution as the reason for this increment (Prem, Vargas and Mejía, 2023).

2.2.2 Colombia's education system

Colombia's education system is regulated centrally by the Ministry of Education (MEN), but it is administered 'locally' by local certified entities, or ETC,[8] in accordance with MEN'S regulations. All children between five and fifteen years old are legally required to attend one year of preschool, 5 years of primary school, and 4 years of middle school (lower secondary education), the only compulsory levels (MEN, 2013), yet from 7 to 13 years is the age range at which over 90 percent of the school age population is enrolled (OECD, 2016). In order to access tertiary education, students must finish the 2 years of high school (upper secondary education) and take the SABER 11 standardized test.[9] Children usually enroll into preschool (grade 0) at age 5 and require 12 consecutive years of schooling to finish high school, at age 17.

School attainment varies depending on the region. The average years of schooling in rural

[8]Colombia is administratively divided into 32 *departamentos* and Bogota, the capital district. These are equivalent to the US states. The departamentos are divided into *municipalities*, the smallest administrative unit. There are 96 *entidades territoriales certificadas* (ETC) and they can be either a departamento, a municipio, or a district.

[9]The SABER 11 test is administered prior to graduation in Colombian high schools (grade 11th). It is similar to the SAT in the United States, but it is administered by the Colombian government.

areas is 5.5 years while that of its urban counterpart is 9.6. In 2014, 87% of schools were public and 78% were located in rural areas (OECD, 2016). Areas where there is presence of coca crops present relative lower values or coverage and quality than other parts of the country, as well as higher poverty levels (Zuleta, 2019). Extremely poor farmers are more likely to grow coca than non-poor farmers in the same areas (Dávalos and Dávalos, 2020). In these areas, it is not uncommon for children to join the labor force (Rodríguez, 2020) or for them to skip school during harvesting season.

2.3 Data and Research Design

2.3.1 Data

I built a municipality-year level panel data to study the effect of coca presence on school participation. In particular, I use the increasing area affected with coca crops to study its effect on school dropout, enrollment, and retention. I focus on the period from 2012 to 2016, during the President Santos term only, who initiated peace negotiations with FARC guerrilla in 2012. Notice that the period of the peace process overlaps with that of constant growth of the area effected with coca bushes, and contains the year of the suspension of the aerial spraying strategy (see Figure 2.1).My sample consist of all 1,120 municipalities of Colombia, except for the 2 municipalities that form the departamento of San Andrés, Providencia and Santa Catalina and are not in the mainland.[10] I describe the main variables and the data sources below.

[10]The state of San Andrés, Providencia and Santa Catalina is located in the Caribbean Sea, approximately 482 miles northwest of mainland Colombia. I drop them from the sample due to their location and because the presence of coca crops is not monitored.

2.3.1.1 Education data

The main dependent variables are school-related outcomes that come from the MEN. This data has information aggregated at the municipal level each year over the period 2012 to 2016. It has the enrollment and dropout rates, as well as the rate of students that approved, failed, and repeated primary, middle and high school, separately. Note that, students enrolled either pass, fail, or dropout, and among those that fail, some students repeat the grade while others also dropout or change of school. My analysis is focused on the enrollment, dropout, and retention rates. A limitation, however, is that it does not have an attendance variable that captures short term dynamics of students. It also does not have information on either households' or students' migration decisions to study. It is important to clarify that, even though the outcome comes from a decision from the individual, my treatment and data is at the municipal level.

The dropout rate captures yearly dynamics as it is measured as the share of students that leave school at any point during the academic year relative to the initial enrollment. Under this methodology, for instance, a student that changes school during the academic year will be counted as if she has dropped out of school. On the other hand, the enrollment rate is measured as the number of students of an age group that enrolled in school at the beginning of the academic year relative to the size of the population of that age group (MEN, 2013). Notice that the latter rate can take values greater than one because of overage students outside the reference age group or student attending school from neighboring municipality. It is not possible to distinguish between these two however, the latter is more likely to happen in more urban municipalities where students from surrounding areas come to study, while the former is more likely to happen in the more rural municipalities as the access to school is more limited (Forero, 2019). The grade pass and grade fail rates capture the share of students that fulfill the academic requirements of their grade, and the ones that do not fulfill them, respectively. These two rates are aggregated into the retention rate that speaks of the share of students that did not drop out from school. The grade repetition rate reports the share of students that have to repeat or go through an academic grade again (MEN, 2013).

2.3.1.2 Illegal crops and spraying data

Data on illicit crops comes from the Integrated Monitoring System of Illicit Crops (SIMCI by its Spanish acronym) of the United Nations Office on Drugs and Crime (UNODC). It has information on the municipal area (measured in hectares) cultivated with coca bushes each year between 2001 and 2018.[11] SIMCI is a satellite-based monitoring system that uses imagery of the entire territory of Colombia's mainland to estimate the extension of the coca crops annually, as of December 31 of each year. 70 percent of the satellite images are obtained within two months range (two months before or after the cut-off date), while of the remaining 30 percent, about half is obtained from August to November of the estimate year, and half between March and April of the following year (Abadie et al., 2015). Coca plots identification are based on visual interpretation of satellite images, including three stages: (1) preliminary visual interpretation includes an analysis of the historical coca series and georeferenced photographs taken in overflights by the National Police, (2) verification overflights based on visual inspection from an aircraft on the territories affected by coca crops, and (3) edition using the verification overflights information to adjust the preliminary interpretation (UNODC, 2017). Note that, using this methodology, an area that was affected during a particular year can be reported as not having coca if eradication was

[11] Data availability goes back to 1999, although starting 2001 the data evaluates the entirety of the country.

effective in the territory.[12]

Reports on illicit crops eradicated using aerial spraying are produced by DIRAN. This information is available from 1999 to 2015, when aerial spraying was suspended. The aerial spraying of coca crops was the most important strategy during the first years of *Plan Colombia* and its use declined in the last years. Figure 2.1 presents evidence of this by showing the area eradicated by aerial spraying present a downward trend from 2006 until its suspension. There is no information on what municipalities were sprayed or how many times, however following the previous discussion on effectiveness, I assume that all the municipalities that report positive values of eradication in a given year were sprayed in that year. Using this definition, it exists the possibility of labeling some areas as "non-sprayed" where in fact those areas were sprayed. However, had this be the case, the impact of the intervention would be effective.

2.3.2 Evaluation design

The ideal experiment would compare school outcomes in municipalities that are identical in all respects except exposure to coca fields. To approximate such an experimental ideal, I use municipality-level variation on the share of area affected with coca fields. However, the municipalities that have presence of coca bushes could be systematically different the ones that do not have such presence leading to a selection bias problem.

I take advantage of the increasing trend in coca cultivation and use the exposure to fumigation missions as a source of exogenous variation, as not all the areas affected with coca presence are sprayed. The targeted areas are dynamically determined by changes over time and over mu-

[12] Abadie et al. (2015) discuss the advantages of using this methodology relative to alternative ways of measuring coca cultivation at the municipal level.

nicipalities, although they present a persistent location pattern (UNODC, 2017, 2019). Figure 2.3 and Figure 2.4 show this pattern using 3-years of data. The removal of the aerial spraying missions can be seen in panel (d) of Figure 2.4 where any of the municipalities is shown to be sprayed, even though there were still areas affected with coca. Notice that the areas affected are persistent in time, but it is important to highlight that this figures do not allow me to separate whether the existing fields affected were increasing in plot size or moving within the municipality.[13]

I exploit the temporal variation created by the temporary effects of the fumigation on the coca fields to instrument the presence of coca. I also use the cross-sectional variation of the fumigation strategy as not all the municipalities affected with coca presence are effectively sprayed. Notice, spraying missions are administered and centrally planned, then I can hardly claim that fumigation is random; however, if the missions take place in an area where there is coca presence and all areas are equally accessible, then I can test whether the probability of spraying is given by the presence of coca crops and by geographic characteristics.

To determine the probability of a municipality being sprayed, I use geographic characteristics at the municipal level that help explain the presence of coca crops (Acevedo and Bornacelly, 2014; Mejía and Restrepo, 2015), and information on the area cultivated. The former set of characteristics do not vary over time while the share of area cultivated does vary. Assuming, for simplicity, that I define an area as fumigated as

$$D_{mt} = \mathbb{1}\{Spray_{mt} > 0\}$$

[13] Prem, Vargas and Mejía (2023) findings suggest that the increase in coca fields comes from a growth in pre-existing plots as well as in plots detached from pre-existing ones.

where $\mathbb{1}\{\cdot\}$ is the indicator function and D is a dummy variable that takes the value of 1 when the municipality m in year t reports a positive area eradicated through an aerial spraying mission, $Spray_{mt}$; and 0 otherwise. Then, I estimate the following linear probability model

$$D_{mt} = \beta coca_{mt} + X_m \delta + e_{mt} \ , \tag{2.1}$$

where $coca_{mt}$ is a measure of coca presence in municipality m and year t normalized by the area of the municipality, and X_m is a vector of geographical characteristics that do not vary over time. Specifically, it includes the distance to the closest airport from where spraying missions depart, the number of airports within the 80 miles radius, the distance to the Department's capital and closest market, the share of rural population and area, and the characteristics used to create the coca suitability index (Mejía and Restrepo, 2015).

Table 2.1 report the estimates. Column (1) presents the results using the full sample of municipalities, including those without presence of coca crops. The estimates suggest a positive relationship between the presence of coca crops and the probability of spraying, as well as a significant relationship with some of the geographic characteristics. However, further examination of the data revealed that this result is driven mainly by the municipalities that do not have coca crops and are not sprayed. In Column (2), I present the estimates restricting the sample to the subset of municipalities that have been sprayed at least one time, i.e., ruling out the ones that either do not have presence of coca or have never been sprayed for other reasons. Interestingly, the results suggests that the share of area cultivated marginally and negatively explains the probability of the spraying mission. Furthermore, some of the geographic conditions explain this probability. For instance, the distance to the principal market and to the Department's capital is positively estimated as expected as coca is grown in more remote areas. In addition, water disposability—which captures areas with more precipitation—is also positively correlated with the probability of spraying. This result is less expected as spraying missions need to have 'optimal' weather conditions that includes less rain so as the glyphosate is not wash out from the crops, but on the other hand, more rains could affect positively the production of the crop (Rodríguez, 2020). Finally, the soil erosion and elevation are negatively correlated with the probability of being sprayed, although these two characteristics are related to the presence of coca crops (Mejía and Restrepo, 2015).

Columns (3), (4), and (5) further explore the dynamics of the spraying campaigns on the municipalities that have been sprayed. Since spraying campaigns could be decided the prior year based on the information available at the time, the results of Column (3) include the lagged value of the cultivated area while those of Column (4) include both the contemporaneous and the lagged values of the cultivated area. These two results go in line suggesting no relation. Finally, in Column (5) instead of the current and the lagged value, I include the variation between to two as the government could focus on the areas that present a higher increase. Similarly to previous results, there seems to be no relation between the presence of coca and the spraying campaigns except for the growth in area cultivated that presents a negative relation but marginally significant. I take these results as suggesting evidence that the decision to spray is quasi-randomly determined based on observable geographic characteristics and not decided deterministically.

2.4 Empirical strategy

To uncover the relation between coca presence and school participation, I am interested in the following relation:

$$school_{mt} = \delta_m + \delta_t + \beta coca_{mt} + \varepsilon_{mt} \tag{2.2}$$

where the parameter of interest is β. However, it is likely that the estimates will be biased given that the presence of coca fields is not random. To tackle this issue, I use the fumigation missions of coca fields as a source of exogenous variation. Particularly, I label as treated a municipality being fumigated (D_{mt}=1) if there is any positive area reported as eradicated through a fumigation mission (see previous section). Formally, this relation can be written as:

$$coca_{mt} = \gamma_m + \gamma_t + \theta D_{mt} + \upsilon_{mt} \tag{2.3}$$

which lead to a difference-in-differences framework. However, my setting differs from that of the 'classical' difference-in-differences as D_{mt} varies within municipalities over time. I use the Fuzzy Difference-in-Differences methodology developed by de Chaisemartin and D'Haultfœuille (2018), which allows me to consider the 'fuzzy' setting by comparing groups who adopt and remove the treatment while controlling for the timing of the fumigation missions, and to recover the Wald estimand of the difference-in-difference.

This comparison deal with (i) the plausibly staggered adoption of the treatment, as there were some municipalities that were first fumigated in different years. (ii) It relaxes the assumption that treatment is an absorbing state, i.e., units exposed to treatment remain treated thereafter, as fumigation has no permanent effect on the coca fields; an area sprayed takes six to eight months to regenerate to a point in which coca can be grown again. This could imply that municipalities switch from being sprayed one year to not sprayed the next one (or vice versa); moreover, (iii) part of the fuzzy design comes because not all the areas affected with coca presence are being sprayed every year and some areas being sprayed report zero area affected due to the effectiveness of the fumigation, i.e. they report control areas as being treated.

The Fuzzy Difference-in-Differences allow me to recover the Wald-DID estimand of the switchers based on pair-wise comparisons of the switchers, i.e., those who adopt the treatment and those who have the treatment removed. The LATE for the "treatment group switcher" is defined by

$$\Delta = E(Y_{11}(1) - Y_{11}(0)|S)$$

where S represent the group of switcher units going from non-treatment to treatment between two consecutive periods. $Y_{gt}(D)$ denote the potential outcomes of Y in treatment group g and period t under treatment D. Similarly, it is possible to define the LATE, Δ', for the control group switchers S', i.e., the switcher units that go from being treated to being untreated between two consecutive periods. Then, the Wald-DID is equal to a weighted average of the LATEs of treatment and control group switchers defined by

$$W_{DID} = \alpha\Delta + (1-\alpha)\Delta' \, , \tag{2.4}$$

where $\alpha = (P(D_{11} = 1) - P(D_{10} = 1))/DID_D$ and $1-\alpha$ are the weights, and DID_D is defined as $E(D_{11}) - E(D_{10}) - (E(D_{01}) - E(D_{00}))$. Notice, the Wald-DID allows for the possibility of heterogeneous treatment effects between groups, that is, $\Delta \neq \Delta'$.

To recover the Wald estimands of my parameter of interest β, I use the stata package `fuzzydid` provided by the same authors (de Chaisemartin, D'Haultfœuille and Guyonvarch, 2019). Standard errors are block bootstrapped at the municipal level in all regressions (Abadie et al., 2022).

Intuitively, I expect the suspension of the aerial spraying to reduce the uncertainty of coca growing farmers and then increase the production of coca crops. At its turn, this increase in production should lead to an increase in the demand for unskilled labor force. This excess in demand is accompanied by an increase in the labor supply through more children joining the labor force, especially from the upper grade level. This implies that I expect a negative sign for the enrollment and retention rates and a positive sign for the dropout rates.

2.5 Results

2.5.1 Coca presence and school outcomes

Table 2.2 presents the Wald estimate for the main outcome variables using the Fuzzy Difference-in-Differences strategy. Each column of the table uses a different outcome as reported in the column header. The outcome variables are standardized to have zero mean and unit variance in the sample. As expected, the estimates for dropout rate for the three school levels are positive and for enrollment rates are negative, although only the high school dropout rate is statistically different from zero at the 5 percent level. This result goes in line with the idea that teenage students are more likely to join the labor force in areas sprayed (Rodríguez, 2020). In particular, a one percent increase in the area affected with coca crops cause an increase in 0.3

standard deviations in high-school dropout rate. This implies that, on average, the high school dropout out rate increase by almost 0.9 percentage points. Given that the average rate is 3.15, this results suggest an effect of an increase in almost 30%.

On the other hand, the results in the retention rate are the complement to those of dropout rate but they need not to be exact opposites. As in the case of the dropout rate, both retention rates for primary and middle school are not statistically different from zero, while the one for high school is statistically significant at the five percent level. In terms of size, they are much smaller with only a decrease of about 0.1 percentage point, however given that the students that remain in school are the ones that either fail or pass the academic year, the results are less obvious. I further explore these results in section 2.5.2.

2.5.2 Students retained

In this section, I explore what happens with the students that do not drop out from school. It is comprised of students who do not dropout from school but finish the academic year either fulfilling the academic requirement of their grade (pass) or not fulfilling them (fail). Finally, among those who do not fulfill the requirements, there exist those who have to repeat the grade for academic reasons. This is a concern in Colombia's education system as the proportion of grade repeaters of 15-years-old having to repeat at least one year is 41 percent (OECD, 2016). A possible explanation for the low academic performance of students is that they skip school, for instance, during harvesting time. Note that, had the students decide to skip school, their learning would be affected. Therefore, the share of students who fulfill the academic requirements will decrease, and the share of students who do not fulfill them will increase. The expected sign for the grade repetition rate is not obvious because students who fail a grade might have incentives to retire themselves from the education system, and not to repeat the academic grade. However, if this is not the case, the expected increase in the share of students that fail might lead to more students repeating a grade.

I follow the same empirical strategy using as outcomes the share of students that fulfill, do not fulfill the academic requirement, and those who repeat. For simplicity, I will refer to these grade pass, grade fail, and grade repetition rates, respectively, and report the results in Table 2.3. All the estimates are not statistically different from zero, except for the high school grade pass rate that present a negative sign and significance only at the 10 percent level. This result is expected for primary and middle school as the presence of coca fields seems to not be related to students leaving school. In addition, the students that are not dropping out are performing worse than their counterparts in places without coca presence. Therefore, this result goes in line with the hypothesis that students miss school during harvesting seasons, although they do not drop out of school completely. Yet, there is no finer data to test this hypothesis.

2.5.3 Test scores

To further explore what happens with the students that remain in school, I use data on standardized test scores to see whether there is an effect on their learning outcomes. Data on these tests comes from the ICFES, a Colombian governmental agency in charge of measuring the quality of education. I am using information on the high school exit standardized test (namely, Saber 11), and from the end of primary and middle school test (Saber 5 and Saber 9, respectively). These exams are used to measure the quality of the education institutions and the high school exit exam is also used by the students to access higher education. The data cover the period 2012 to 2016, however due to comparability of test, the year 2012 and 2013 results for Saber 11 are excluded from the sample. Outcomes are standardized to have zero mean and unit variance in the sample.

I use the same strategy using the test scores averaged at the municipal level as outcome and present the results in Table 2.4 for the mathematics and language components of the Saber 5, 9, and 11 exams. Surprisingly, most of the estimates are not statistically different than zero, except for the mathematic results in primary school students. The results suggests that students that stayed in primary school perform better. This idea goes in line with the opportunity cost of students staying in school and with heterogeneity of unobserved preferences. That is, students that decide to stay in school are those whose opportunity cost of leave is higher because they are potentially better students, but also, they are receiving a more personalized education.[14] Nonetheless, given that there are no more students dropping out of school, as the previous results showed, this is somewhat surprising. The rest of the coefficients present a positive point estimate, which is consistent with the hypothesis, except for the high school language exam. The point estimate does not have the expected sign, although it is not statistically different than zero. Note that these results might be driven by other factors like the presence of civil conflict, as suggested by Rodríguez and Sánchez (2010). The relation between presence of coca fields and civil conflict has been proven in previous studies, as well as the relation between civil conflict and school attainment. However, if this were the case, then the presence of both coca crops and civil conflict should have a bigger impact on test scores as both effects go in the same direction.

[14] Vargas, Gamboa and García (2014) found a similar result using the municipalities in Colombia that faced civil conflict.

2.5.4 Legal crops

So far, my results suggests that there are larger dropout rates among school-age individuals in the areas with more coca fields, in particular for high-school students. Similarly, these students are also not fulfilling the requirement to advance to their subsequent grade yet, those that finish high school have not differential effects on their exit exam. There is no evidence that students are dropping out from school to join the labor market or coca related activities. Since school-age students may have more difficulties joining the formal labor market, I can explore whether there are changes in the production of other agricultural crops.

I shed light on this mechanism by using data from the Ministry of Agriculture, gathered by Universidad de los Andes (Acevedo and Bornacelly, 2014), that report for a variety of agricultural products the area devoted to cultivation and the area harvested. Using these two measures, I create the ratio of area cultivated to area harvested to explore the possibility that there are less crops being harvested relative to what is cultivated. In addition, I calculate the share of area cultivated and harvested relative to the municipality area to explore the possibility independently as the presence of coca crops could be affecting both variables. These measures will allow me to evaluate whether there is a crowd out effect of the coca crops on the legal crops.

My results suggests that the presence of coca crops have no effect on the legal crops, neither in the area cultivated nor in the area harvested (Table 2.5). These results suggest the absence of a crowd out effect due to the presence of coca crops. One could interpret these results as the agricultural labor market being in equilibrium despite the coca surge, i.e., the coca crops are not replacing the licit crops and the labor force is fully allocated. If this were the case, it opens the possibility to school-age students to fill the excess in labor demand on coca-related activities.

Overall, these results go in line with the hypothesis that school-age students are joining the labor force to work on coca related activities (Rodríguez, 2020; Sviatschi, 2022). However, I cannot claim that this is in fact the case, as I do not have data to support it, but I cannot rule it out either.

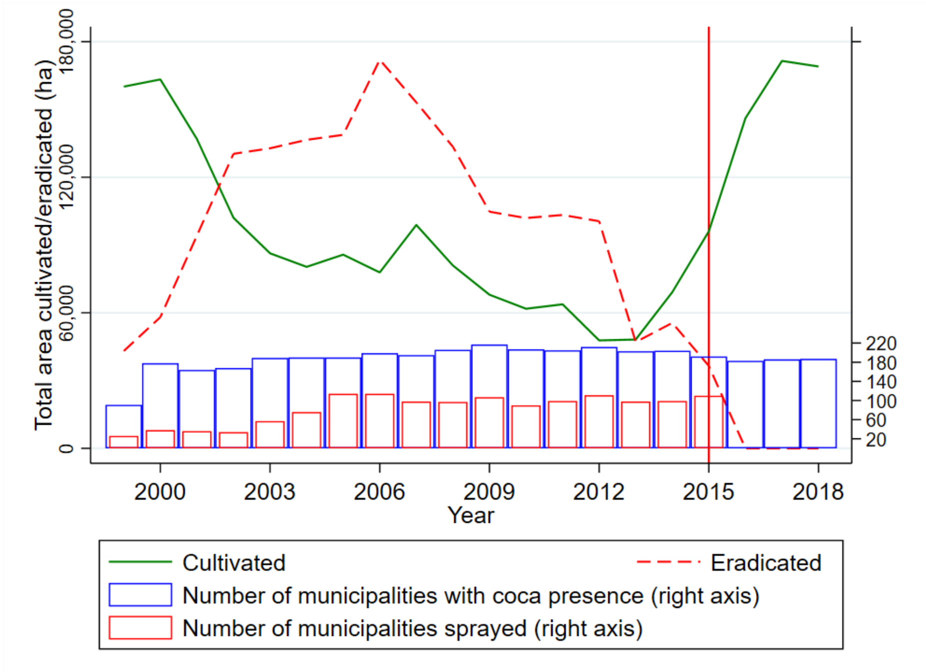
2.6 Conclusion

I provide evidence that school-age individuals exposed to agricultural shocks experience less human capital accumulation. I focus on the case of the illegal market of coca leaves, the primary input in cocaine production. I use the case of Colombia as it carry out fumigation missions where coca fields are aerially sprayed with herbicide as part strategy to force eradicate coca crops. I exploit no permanent effect of this fumigation missions to test my hypothesis. This temporal effect led me to an instrumental variable difference-in-differences setting. However, since not all the assumptions were met, I borrow from the literature the Fuzzy Difference-in-Differences methodology to tackle the challenges in the assumptions.

I find that the increasing presence of coca crops is associated with more students dropping out of high school, while I do not find any effect on the other grades nor in the enrollment rate. My results also suggests that high school students are less likely to fulfill the requirement to advance to the next grade. Finally, I explore possible effects on learning outcomes and find that only primary school students mathematics test score is positively affected, a result consistent with previous studies. I rule out the possibility that school-age students are dropping school for working in legal agricultural activities which—along the other results— I interpret as suggestive evidence that school-age students that drop out from school are working in coca related activities.

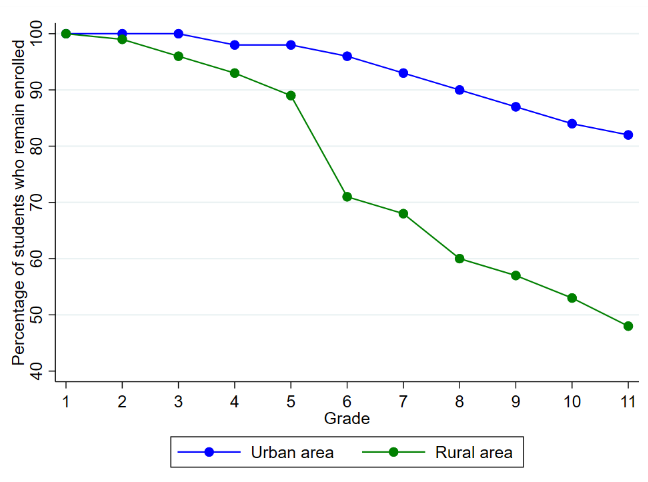
Overall , this paper provide an additional step at understanding the role of illegal markets in schooling decisions and schooling outcomes in the presence of external shocks.

Figure 2.1: Dynamics of the coca cultivation and eradication in Colombia, 1999-2018



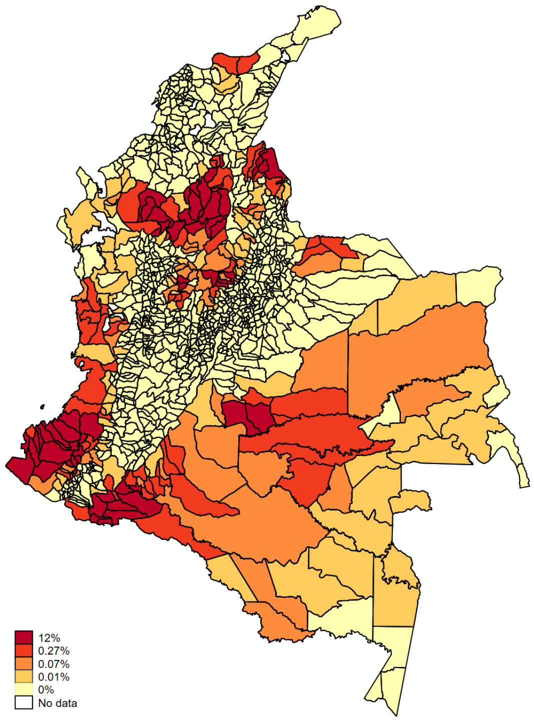
Note: Data from UNODC-SIMCI and DIRAN. Over the period, there has been 327 municipalities that have had presence of coca at least in one year, and 235 that have been sprayed at least once. There are 208 that have had presence of coca and have been sprayed at least once. The total area in hectares is reported in the left axis, while the number of municipalities affected and sprayed is reported in the right axis.

Figure 2.2: Percentage of students who remain enrolled by grade

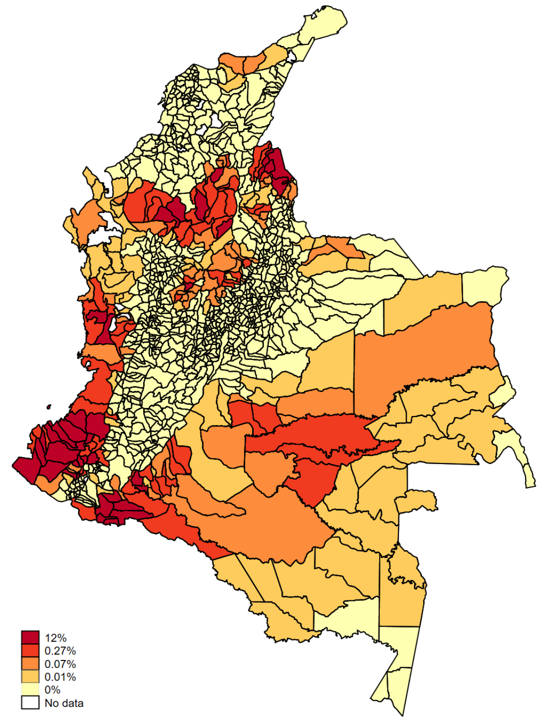


Source: Taken from MEN (n.d.). Data from 2008 Quality of Life Survey.

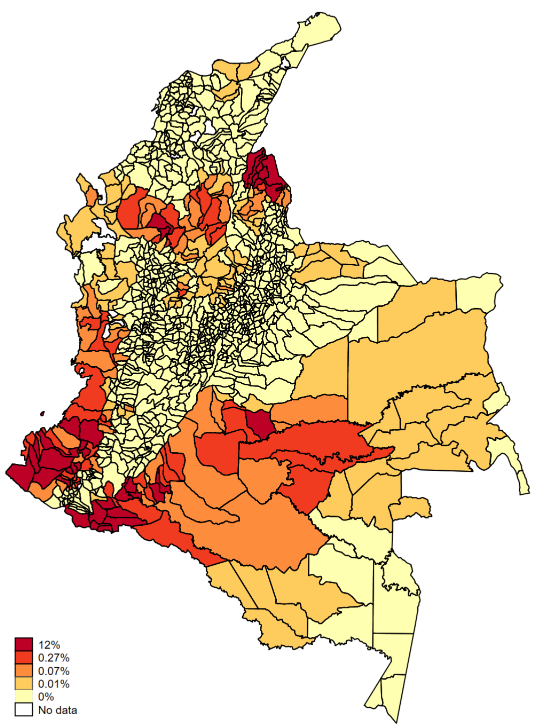
Figure 2.3: Areas affected by coca presence, 2007-2018



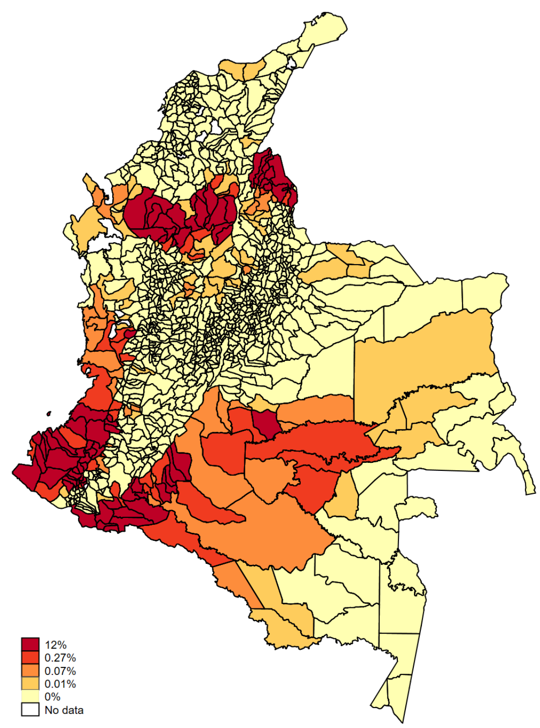
(a) 2007 to 2009



(b) 2010 to 2012



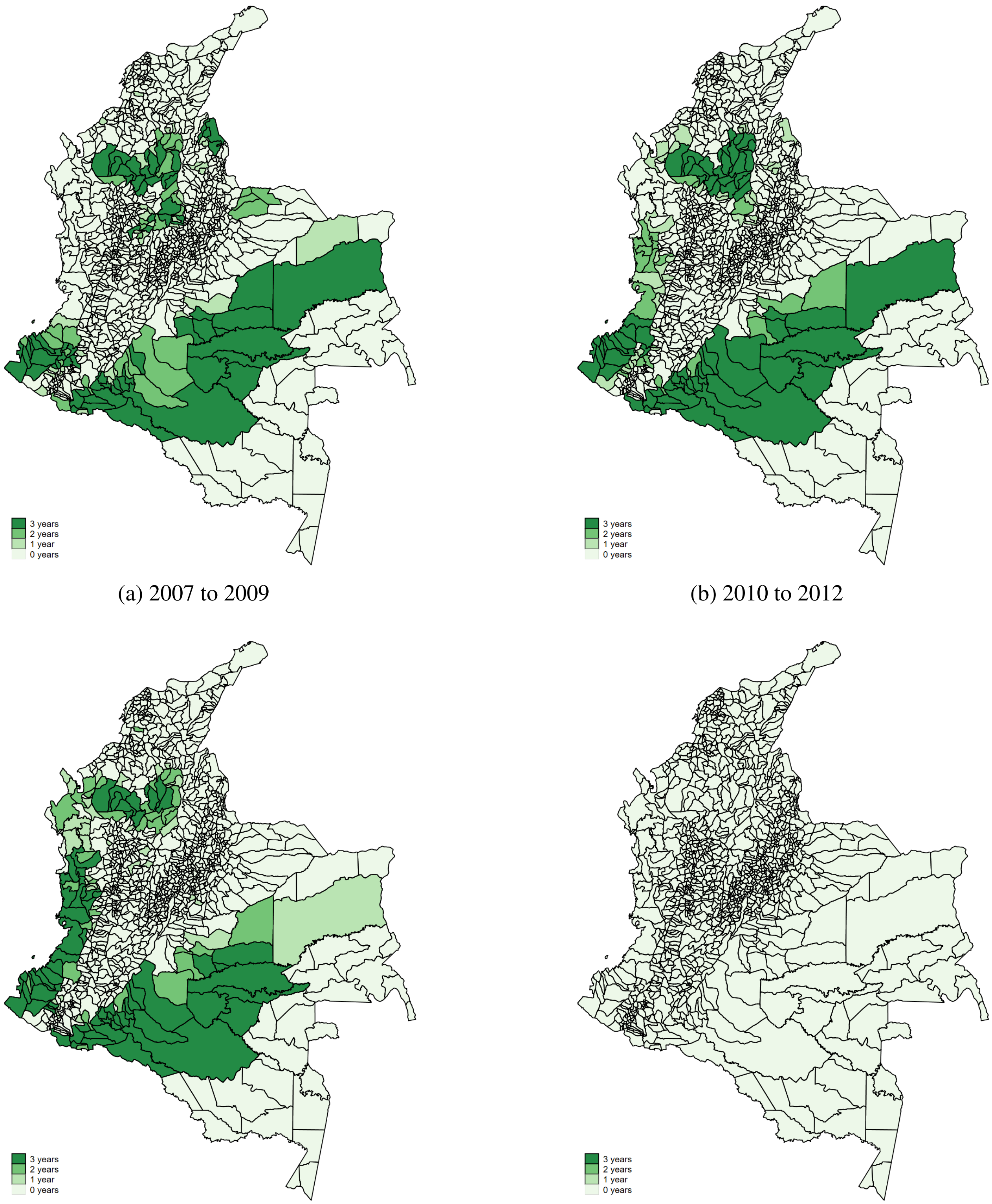
(c) 2013 to 2015



(d) 2016 to 2018

Note: Data from UNODC-SIMCI. The area affected is measured as the share of area with presence of coca crops, relative to the total area of the municipality. Panels (a) to (d) show the three-years average of the area affected from 2007 to 2018 using the same scale. Darker areas represent more affected municipalities.

Figure 2.4: Territories sprayed, 2007-2018



(a) 2007 to 2009

(b) 2010 to 2012

(c) 2013 to 2015

(d) 2016 to 2018

Note: Data from DIRAN. A municipality is considered sprayed if the area eradicated by aerial spraying is positive. Panels (a) to (d) show the three-years sum of the area affected from 2007 to 2018. Darker areas represent municipalities that were sprayed more times.

Table 2.1: Probability of sparying

VARIABLES	(1) Prob=1	(2) Prob=1	(3) Prob=1	(4) Prob=1	(5) Prob=1
Share of cultivated area	0.005***	-0.002*		-0.007	
	(0.001)	(0.001)		(0.005)	
Lag. Share of cultivated area			-0.002	0.007	
			(0.002)	(0.006)	
Δ Share of cultivated area					-0.007*
					(0.004)
Number of airports in radius	-0.002	0.030	0.030	0.030	0.030
	(0.009)	(0.034)	(0.033)	(0.034)	(0.034)
Distance to closest fumigation airport	-0.056***	-0.010	-0.009	-0.009	-0.009
	(0.014)	(0.034)	(0.034)	(0.034)	(0.033)
Distance to Department's capital	0.009*	0.022*	0.021*	0.021*	0.021*
	(0.005)	(0.013)	(0.013)	(0.013)	(0.013)
Distance to closest principal market	0.094***	0.218***	0.215***	0.216***	0.216***
	(0.016)	(0.052)	(0.052)	(0.052)	(0.050)
Elevation x 100 mts (MASL)	-0.002***	-0.006*	-0.006*	-0.006*	-0.006*
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
Water disposability index	0.045***	0.100***	0.100***	0.100***	0.100***
	(0.006)	(0.020)	(0.020)	(0.020)	(0.020)
Soil erosion index	-0.016***	-0.044**	-0.044**	-0.043**	-0.043**
	(0.004)	(0.020)	(0.020)	(0.020)	(0.020)
Soil aptitud index	-0.018***	-0.017	-0.016	-0.017	-0.017
	(0.005)	(0.019)	(0.019)	(0.019)	(0.019)
Share of rural population	0.003	-0.036	-0.037	-0.037	-0.037
	(0.022)	(0.096)	(0.096)	(0.096)	(0.096)
Observations	6,348	1,308	1,308	1,308	1,308
R-squared	0.125	0.132	0.131	0.133	0.133

Notes: This table presents the probability of a municipality being sprayed. Column (1) include all the municipalities while Column (2) to (5) results restrict the sample to municipalities that have been sprayed at least once since 2001. Standard errors reported are clustered at the municipality level and presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level. .

Table 2.2: Main results: Effects of coca presence in school enrollment and dropout

Dependent variable:	Dropout rate			Enrollment rate			Retention rate		
	Primary	Middle	High	Primary	Middle	High	Primary	Middle	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Area affected coca	0.109	0.136	0.328**	-0.028	-0.004	-0.022	-0.118	-0.104	-0.259**
	(0.133)	(0.128)	(0.157)	(0.021)	(0.032)	(0.046)	(0.124)	(0.169)	(0.105)
Observations	7,653	7,627	7,520	5,521	5,521	5,521	7,653	7,631	7,570
Dep. Var. Mean	2.73	4.57	3.15	86.93	90.15	73.35	97.27	95.43	96.84
Dep. Var. S.D.	2.12	3.20	2.71	22.78	23.81	22.78	2.09	3.18	2.70

Notes: Estimates in this table computed using the Stata command `fuzzydid` (de Chaisemartin, D'Haultfœuille and Guyonvarch, 2019). Each entry in the table comes from a different estimation. Outcomes are reported in the header of the column and each one is standardized to have zero mean and unit variance in the sample. Standard errors are clustered at the municipality level and bootstrapped using 50 repetitions. They are reported in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 2.3: Effects of coca presence in schooling outcomes

Dependent variable:	Grade pass rate			Grade fail rate			Grade repetition rate		
	Primary	Middle	High	Primary	Middle	High	Primary	Middle	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Predicted Area affected coca	0.098	0.024	-0.179*	-0.234	-0.152	0.026	0.126	0.027	-0.178
	(0.127)	(0.125)	(0.095)	(0.167)	(0.107)	(0.145)	(0.115)	(0.114)	(0.128)
Observations	7,721	7,699	7,661	7,653	7,648	7,618	7,599	7,595	
Dep. Var. mean	94.01	89.97	93.46	3.11	5.28	3.18	1.72	2.36	0.97
Dep. Var. S.D.	4.58	7.08	5.13	3.40	5.78	3.73	2.01	2.72	1.43

Notes: Estimates in this table computed using the Stata command `fuzzydid` (de Chaisemartin, D'Haultfœuille and Guyonvarch, 2019). Each entry in the table comes from a different estimation. Outcomes are reported in the header of the column and each one is standardized to have zero mean and unit variance in the sample. Standard errors are clustered at the municipality level and bootstrapped using 50 repetitions. They are reported in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 2.4: Effects of coca presence on test scores

Dependent variable:	Language Primary Saber5 (1)	Language Middle Saber9 (2)	Language High Saber11 (3)	Mathematics Primary Saber5 (4)	Mathematics Middle Saber9 (5)	Mathematics High Saber11 (6)
Area affected coca	0.143	0.121	-0.036	0.174**	0.072	0.012
	(0.089)	(0.082)	(0.034)	(0.086)	(0.085)	(0.039)
Observations	6,555	6,536	4,366	6,553	6,534	4,366
Dep. Var. mean	295.51	284.33	48.78	294.78	290.51	47.86
Dep. Var. S.D.	30.79	31.95	3.68	33.30	33.8	4.28

Notes: Estimates in this table computed using the Stata command `fuzzydid` (de Chaisemartin, D'Haultfœuille and Guyonvarch, 2019). Each entry in the table comes from a different estimation. Outcomes are reported in the header of the column and each one is standardized to have zero mean and unit variance in the sample. For primary and middle school, estimates uses data from 2012 to 2016, while for high school, from 2014 to 2016. Standard errors are clustered at the municipality level and bootstrapped using 50 repetitions. They are reported in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 2.5: Legal crops and coca presence

Dependent variable:	Ratio harvest-cultivated (1)	Share area cultivated (2)	Share area harvested (3)
Area affected coca	0.004	0.004	0.005
	(0.012)	(0.004)	(0.005)
Observations	5,403	5,459	5,459
Dep. Var. mean	0.850	0.141	0.121
Dev. Var. S.D.	0.124	0.166	0.145

Notes: Estimates in this table computed using the Stata command `fuzzydid` (de Chaisemartin, D'Haultfœuille and Guyonvarch, 2019). Each entry in the table comes from a different estimation. Outcomes are reported in the header of the column. Standard errors are clustered at the municipality level and bootstrapped using 50 repetitions. They are reported in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Chapter 3: Cheer or Fear: The 2018 Colombian Presidential Election

3.1 Introduction

Voters' political choices can be molded by the presence of migrants. Which platform benefits depends on the composition of the migrants arriving to the country. Historically, right-wing parties, and parties with an anti-immigrant agenda, have benefited from their presence by associating the presence of migrants to more insecurity or higher unemployment (Halla, Wagner and Zweimüller, 2017). However, when immigrants are important for the prospects of the domestic economy (Edo et al., 2019) or when they have few dissimilarities with natives in terms of language, religion, and race (Mendez and Cutillas, 2014), the presence of migrants affects positively the votes for pro-immigrants and left-wing parties. Most of the empirical evidence to date has focused on South-North migration where migrants relocate looking for a gain in welfare irrespective of the political context of their country of origin. Less attention has been given to migration between developing countries, or South-South migration, which is mostly motivated by ecological disasters or civil conflicts. In some of these instances, the political context of the country of origin plays a role in the decision to migrate.

I study how the exposure to migrants influences the preferences of voters. I consider the case of the 2018 presidential election in Colombia where political campaign of the leading right-

[0]This paper received support from the Institute for Humane Studies, grant no. IHS016907.

wing party, *Centro Democrático* (CD), used the presence of migrants to demonize the agenda of their political contenders. The party suggested that a left-wing government will worsen the socio-economic conditions of the country as the left-wing government in power did in Venezuela. It mobilized voters through the fear that Colombia will turn into a "segunda Venezuela" (second Venezuela) or "estado castrochavista" (Castrochavist state) if a left-wing candidate was elected president in Colombia. I build on the idea that this message could have resonated more in areas where there was a larger presence of Venezuelans because of more first-hand exposure of local voters to migrants.

I exploit the unanticipated inflow of Venezuelan migrants into Colombia through the ports along the land border between the two countries. Venezuelans began to migrate to Colombia at a small but stable rate, through both air and land, shortly after Hugo Chavez raised to power in 1999 and instated a left-wing government. Migration to Colombia through land was suspended between September 2015 and August 2016 due to a diplomatic crisis. After entry ports reopened, Colombia experiences a sharp and unanticipated increase in the inflow of migrants fleeing from the worsening economic condition in their country. For instance, between 2012 and 2015, the share of migrants that enter through land ports represented around 35% of the total migration. This share increased to more than 75% after 2016 when the number of migrants increase more than five times.

The ideal experiment would compare electoral outcomes in the 2018 elections across municipalities of Colombia with as good as random variation in the inflow of Venezuelan migrants after the ports reopened. I take this idea to the data by seeking exogenous factors that may have triggered such random variation. I construct an instrumental variable using the distance of each municipality to the closest land entry port to instrument for the presence of Venezuelan migrants.

I build on the idea that most Venezuelans entered through one of the five border crossings and sorted more disproportionately in municipalities closer to the border.

My identification strategy relies on the assumption that the distance of a municipality to the border predicts the presence of migrants, and it is not correlated directly with the electoral results of the 2018 elections and the political preferences in said municipality. Notice, this is likely to be the case as local support is based on a clientelist system of political exchange with few differences in the socio-economic origins of supporters of different parties (Fergusson et al., 2021). Nonetheless, I provide some evidence of these political preferences using previous presidential election results.

I use data on electoral results at the municipal level in the 2018 presidential election. I observe the share of votes for each candidate in every municipality and voters' turnout. I classify candidates based on their ideology, or that of the political party endorsing them, based on the methodology proposed by Fergusson et al. (2021). I construct an indicator for the presence of Venezuelan migrants in a municipality by using data from the 2018 Colombian population census. Specifically, I consider those respondents: (i) born in Venezuela, (ii) not living in Colombia in 2013 (five years before the census), and (iii) living in Colombia in 2017 (one year before the census). I show that this measure largely captures the change in Venezuelans as a percentage of the municipality's population between the 2018 and the 2005 population censuses, and therefore it is a good approximation for the inflow of Venezuelans that migrated after 2016.

My empirical findings suggest that a one percent change in the presence of migrants in a municipality raised the share of voters for the CD party by 8 percentage points. Notice, this is a sizable effect as the average share of votes for the CD party was 46.9%. This large coefficient is partly due to the low variability on the measure of the presence of Venezuelan migrants (less than

1% standard deviation). Less restrictive measures present larger variability across municipalities, and the estimated coefficient smaller in magnitude, although sign and significance are not affected. This conclusion is robust to the inclusion of a number of controls and different measures of distance.

In addition, I analyze how the presence of migrants affects the votes for each candidate and find that it polarized the election between two candidates. I find that the increase in votes for *Centro Democrático* is at the expense of the votes for the left-wing candidate, while no other right-wing candidate benefited for the presence of Venezuelan migrants. This last finding marks a difference with the previous literature which associated increasing share of votes for right-wing parties with higher presence of migrants. Most of the literature has grouped political parties by ideology or by government coalition to analyze the effects of the presence of migrants on their share of votes without accounting for different dynamics across parties. These aggregated results led to the conclusion that all right-wing candidates benefit from the presence of migrants. However, my results suggest that the presence of migrants has an effect on the votes for right-wing candidates solely through the votes for one of them. On the other hand, consistent with the literature, I find that the increase in presence of migrants is also associated with an increase in the voter turnout.

I examine the potential mechanism driving these effects. I consider whether political preferences may be explained by changes in crime levels caused by the presence of migrants. I do not find a significant effect of robbery or in homicide rates. It is possible that the changes are due to the perception of crime, although I cannot test this hypothesis. I also examine the possibility that political preferences may be caused by political preferences at the local level. For this purpose, I examine the effect of the presence of Venezuelan on the share of votes for a right-, center-, or left-wing candidates in the 2015 mayoral election. I find that the presence of migrants affected positively the share of votes for right wing-candidates at the expense of the votes for the center-oriented candidates, contrary to the findings of (Rozo and Vargas, 2021)

I provide some suggesting evidence in favor of the strategic electoral misinformation (Rozo and Vargas, 2021). I use the number of internet searches for the words "castrochavista" and "segunda Venezuela", two of the most common phrases used by right-wing parties to campaign against candidates ideologically aligned with the Venezuela regime. I show that the use of this phrases peak during the election months in the 2018 election while it did not during the previous presidential election but during the peace plebiscite. I examine whether the political preferences were shaped by the attitude towards the plebiscite to ratify the peace agreement between the government and the FARC left-wing guerrilla. My findings suggest that the presence of Venezuelan migrants increased the share of votes for the No option.

Finally, I study the relative importance of these mechanisms to explain the political preferences of voters in the 2018 presidential election. I examine these mechanisms together and find that the attitudes toward the peace agreement explain about a third of the effect of the presence of Venezuelan migrants on the political preferences of voters in the 2018 presidential election.

I contribute to the literature on the effects of migration on political outcomes by studying the effect of South-South migration on voting behavior. This migration is not motivated by a welfare increase but rather by social distress or civil conflict. Almost 80% of this migration takes place between countries with contiguous borders (Ratha and Shaw, 2007) and, therefore, attitudes towards migrants are less clear. For instance, political parties in developing countries differ from those of developed countries as they do not necessarily have a stance on migration. Previous literature has suggested that large inflow of international migrants is related to an increase support for the right-wing parties—including extreme right and right-wing coalitions—at either the local or the national level. They focus on South-North migration and, in particular, migration to European countries (Barone et al., 2016; Dustmann, Vasiljeva and Damm, 2018; Edo et al., 2019; Gerdes and Wadensjö, 2010; Halla, Wagner and Zweimüller, 2017; Mendez and Cutillas, 2014; Otto and Steinhardt, 2014). One exception for North America is the paper of Mayda, Peri and Steingress (2022) that explores the effect of immigration on the votes for the Republican party in the United States. There are less studies of South-South migration with the notable exception of Rozo and Vargas (2021) who study the effects of migrants on the voting behavior in Colombia.

My set up allow me to differentiate the effects by political party and, thus, I show that not all right-wing parties benefit from the presence of migrants. Most of the previous literature group political parties by factions or by coalitions to follow them over multiple elections. This feature does not allow them to analyze political parties individually, although it allows them to control for unobserved political preferences.

This paper adds on the literature that explores the effects of the current Venezuelan humanitarian crisis and its effects in neighboring Colombia. Venezuela socio-economic conditions have been worsening for the last 20 years leading to Venezuelan moving out of their country. Rozo and Vargas (2021) studied the effect of their cumulative presence of the voting behavior over the elections in the last 20 years. However, the more dramatic episode started after 2016, leading to more than 4 million of Venezuelan moving to other countries including neighboring Colombia. Their sudden presence has affected educational outcomes (Rozo and Vargas, 2020), electoral outcomes (Woldemikael, 2022), violent crimes (Knight and Tribin, 2023), and cooperative attitudes (Lebow, Medina and Coral, 2020). It also led to the government of Colombia to crease a large-scale regularization program that allow them to join the labor market. The effects of this program have also been recently studied (Bahar, Ibáñez and Rozo, 2021; Lebow, 2022; Santamaria, 2020).

My paper also contributes to the literature that studies the consequences of conflict termination. I show that the attitudes towards the 2016 Colombian peace agreement were an important determinant of the votes for the leading right-wing party. The 2018 presidential election was the first election since the peace plebiscite took place. To the best of my knowledge, this is the only study that includes the peace plebiscite results to explain political preferences and voting behavior. Other papers that study the consequences of conflict termination shows that there is an improvement on educational outcomes (Prem, Vargas and Namen, 2023), a change in the presence of illicit crops (Prem, Vargas and Mejía, 2023), and a change in entrepreneurship dynamics (Bernal et al., 2022).

The paper closer to mine is that of Rozo and Vargas (2021) who study the electoral effects of international migration in Colombia. Their results suggests that political participation increases and that voters shift from left- to right-wing ideologies. My paper is a complement, but it is different to theirs in three ways. First, I use data from the Colombian census to locate Venezuelan migrants at the municipal level rather than predicted their presence. Their prediction overestimates the presence of Venezuelan migrants as they use the number of Venezuelan that enter through the entry port but did not differentiate those that only transit through the country to go to other nations in South America. A phenomenon called “los caminantes” or the walkers.

Second, my strategy focus on the sharp and sudden increase in the presence of Venezuelan migrants. I show that the Venezuelan migrants located disproportionately in municipalities close to the border between the two countries, whereas their prediction is based on early settlements. Their implicit assumption is that the spatial distribution of migrants is stable over time, thus, these two measures capture two different effects, i.e., proximity to the border and early settlements, respectively.

Finally, my results suggest that only the leading right-wing party benefited from the presence of migrants, thus, my results are not sensitive to the criteria used to classify parties by ideology. The study of Rozo and Vargas (2021) groups political parties by ideology, however this aggregation make their results sensitive to the classification of parties by ideology. My results are consistent with their results, the presence of Venezuelan migrants increase the share of votes for the right-wing parties. However, for the case of the 2018 presidential election, I show that this effect can be explained solely by the change in votes for the *Centro Democrático* party.

The rest of the paper is organized as follows. Section 3.2 provide some institutional background. Sections 3.3 and 3.4 explain the main sources of data and the research design. Section 3.5 describes the empirical strategy to estimate the effect of the Venezuelan migration on voting preferences and the validity of the strategy. I present the main results in Section 3.6, while Section 3.7 presents robustness checks and Section 3.8 presents additional results. Section 3.9 presents potential mechanisms and, finally, the last section concludes.

3.2 Institutional Background

3.2.1 Colombian Political Context

Colombia is one of the oldest running democracies in Latin America with a multi-party system. The President is elected by direct vote and by majority rule in a run-off election held the same day throughout the country. If no candidate receives more than half of the votes on election day, a run-off election between the two candidates with the most votes in the first round takes place three weeks later. The winning candidate governs for a 4-year period starting August 7th of the same year and for one term only.[1]

The 2018 presidential campaign was polarized between the leading right-wing party, *Centro Democrático*, and the left-wing candidate, Gustavo Petro. One of the key topics of the campaign was the results of the 2016 plebiscite to ratify the peace agreement between the government of Colombia and the FARC, a left-oriented guerrilla group.[2] Despite the unsuccessful results of the plebiscite, after 50.2 percent of the voters voted against the agreement, the Colombian Congress ratified a revised peace agreement two months later, without conducting a second plebiscite.[3] *Centro Democrático* was the leading party rejecting the peace agreement and they focused their presidential campaign on its adoption. They suggested that, on the one hand, the agreement will instate a left-wing government, and on the other, that this kind of government will create similar conditions to those of Venezuela. The primary message of its campaign was fear of socialism and of left-wing governments by suggesting that Colombia will turn into a "segunda Venezuela" (second Venezuela) or a "castrochavista" state.[4] For this, they used the presence of Venezuelan as an example, pointing out the substantial number of migrants fleeing the worsening socio-economic conditions of their country. Figure 3.1 presents some evidence of their campaign strategy using google search trends. It shows how these words were more searched around the 2018 election and peace plebiscite but not in the 2014 election. The campaign end with the election of the former CD senator and party candidate, Iván Duque, as president of Colombia for the

[1]Immediate re-election was allowed until 2014 when Colombia's Congress, under Juan Manuel Santos presidency, approved a constitutional reform banning immediate presidential re-election, established in 2005 under the presidency of Álvaro Uribe.

[2]Weintraub (2015) showed how the 2014 presidential election was marked by the then-ongoing peace negotiation.

[3]There was no law that mandated the peace agreement to be ratified via a plebiscite. Congress could have ratified the original peace agreement without having it submitted to popular vote.

[4]A portmanteau of Fidel Castro and Hugo Chavez. The term was coined and used by the right-oriented political parties to refer to left-oriented ideologies.

2018-2022 term.

3.2.2 The Venzuelan Crisis

The current political and economic crisis in Venezuela can be tracked back to the election of Hugo Chavez as president in December 1998, although the Venezuelan economy had been declining since the sixties (Bello, Blyde and Restuccia, 2011). His election led to institutional changes including the adoption of the Constitution of the Bolivarian Republic of Venezuela in December 1999 that implemented the so-called 21st century socialism. This political regime is characterized by the implementation of populist, socialist programs, nationalizations, and restrictions on private business (Castillo and Reguant, 2017). The oil boom of the 2000s helped the Venezuelan government to fund these programs by using the profits of the state-owned oil and natural gas company of Venezuela (PDVSA). The oil price plunge of 2014-2016 led to the president of Venezuela, Nicolas Maduro, to declare an economic emergency in January 2016. The decrease in government rents worsened the economic and social crisis, including shortage of food and basic necessities, repression of the opposition,[5] violation of human rights by public authorities (Rozo and Vargas, 2021), and to a sustained inflationary pressure (Knight and Tribin, 2023).

The economic and social situation forced many Venezuelans to migrate out and neighboring Colombia was one of the most picked destinations.[6] Consistent with South-South migration pat-

[5] Hsieh et al. (2011) suggest that the Chavez regime was already repressing opposition citizens. They use information from Venezuelans who sign petitions calling for a vote to remove President Chavez from Office in 2003. After Chavez won the recall election in 2004, signers of the petition lost their jobs and experienced a reduction in their earnings.

[6] According to the United Nations, Population division data, in 2019 alone 2,519,780 Venezuelans migrate out of the country from an estimated population of 28 million inhabitants. South America was the recipient of 74 percent of them. The top three recipient countries were Colombia with 41.6 percent; Peru, 23.9 percent; and United States, 10.1 percent. See media coverage from BBC (2018, 2019) and CNN (2019).

terns, most of the migration takes place between countries with contiguous borders. Particularly, the massive outflow of Venezuelan began in August 2016 when Venezuelan president re-opened the border with Colombia after one-year closure due to a conflict along the border (Knight and Tribin, 2023). Figure 3.2 presents evidence of this event using quarterly data from the official migration points in Colombia. It shows the number of Venezuelans entering the country from the first quarter of 2012 to the second quarter of 2019. The sustained inflow of individuals entering through land entry ports started after the re-opening of the border in the third quarter of 2016, while that of the air entry ports remain stable. Notice, it includes information of all Venezuelan migrants regardless of the purpose of the visit or the length of their stay. However, this figure resonates with the numbers from the Population division of the Department of Economic and Social Affairs of the United Nations that estimate more than one million Venezuelan entering Colombian in 2018 alone. These Venezuelan migrants located disproportionately in municipalities closer to the border and the land entry ports (see Figure 3.3a).

3.3 Data

Electoral data

My main outcomes are the share of votes for each candidate and the turnout rate in the first round of the 2018 presidential election. This data comes from Colombia's electoral agency and has information about the number of votes at the candidate level for each municipality.[7] The first round had 6 candidates running for office[8] and 3 of them averaged around 88% of the total

[7]The municipality is the smallest administrative unit of Colombia. It is equivalent to the USA county.

[8]Vivian Morales renounced her candidacy weeks before the election and endorsed Iván Duque. Morales and her VP formula were also on the ballot and got 0.2% of the votes.

votes. The highest voted candidates were Iván Duque with 39.1 percent of the votes, followed by Gustavo Petro with 25.1 percent and by Sergio Fajardo with 23.7 percent.[9] I classify candidates based on their ideology, or that of the political party endorsing them, based on the methodology proposed by Fergusson et al. (2021): Petro is classified a as left-oriented candidate; Duque, Vargas Lleras, and Trujillo as right-oriented; and Fajardo and de la Calle as non-left, non-right. Abusing the terminology, I will refer to the latter group as center-oriented hereafter.[10]

3.3.1 Measure of Venezuelan presence

I construct an indicator for the presence of Venezuelan migrants using data from the 2018 population census,[11] available from Colombia's Statistics Department (DANE for its Spanish acronym). I count the number of individuals that were (i) born in Venezuela, (ii) not living in Colombia in 2013 (five years before the census), and (iii) living in Colombia in 2017 (one year before the census). I normalize this measure by the Colombian population in that municipality.

This measure captures the individuals that enter Colombia starting in the first semester of 2013 until the first semester of 2017. Figure 3.3a shows how migrants located in municipalities closer to the border based on this measure. Figure 3.3b shows that this measure capture the change in Venezuelans as a percentage of the municipality's population between the 2018

[9]The first four candidates averaged around 95 percent of all votes. German Vargas Lleras, the fourth highest, averaged around 7 percent of all the votes and even obtain most of the votes in various municipalities in the first round. Vargas Lleras was vice-president during Santos second term and resigned in March 2017 to run for office as mandated by the law. He is associated with right-oriented ideologies.

[10]Political parties in Colombia has not been classified before by other sources. For instance, the Manifesto Project that classifies political parties' positions for a variety of countries does not have information for Colombia. Notice, for the purpose of this paper, the classification for the candidates running in the election can be done easily as there are just a handful of them.

[11]The 2018 census information was gathered between January and June in two phases. The first phase allows citizens to fill in the census questions on-line and the questionnaire was available from January 9th to March 8th. The second phase consisted of enumerators visiting the dwelling of both the respondents that fill the on-line questionnaire and those who did not. In the former case the enumerators only ask for the confirmation number and in the latter, they conduct the interview. This phase covered a span of 3 months, from April to June.

and the 2005 population censuses. Reassuringly, the *Encuesta Pulso de la Migración*, a recent survey representative of the Venezuelan population, in its first wave shows that only 2% of the respondents claimed to have entered Colombia on or before 2015, while 5% answered 2016, 20% answered 2017, and almost 34% answered 2018. Furthermore, the measure of all the Venezuelan migrants that enter Colombia starting in the first semester 2013 is almost the same as the change of Venezuelan between 2005 and 2018. Taken together, I take this as suggestive evidence that my measure of presence of Venezuelan migrants is a good approximation to the inflow of Venezuelans that migrated after 2016.

This measure of presence of Venezuelan migrants is not free of measurement error. First, my measure leaves out many migrants that arrived within the last year before the census and closer to the election. Since there is an overlap between the dates of the election and the dates that the census was conducted, I decided not to include those respondents, so my measure does not include individuals that arrived after the election. Second, the census could undercount the number of Venezuelan migrants in each municipality as they could have not been interviewed. Although this is a concern, I do not expect differences in the probability of being interviewed by location. In section 3.7, I show that my results are robust to difference measures of presence.

3.4 Research Design

I am interested on whether the presence of Venezuelan migrants influenced the preferences of voters in the 2018 presidential election. I exploit the unanticipated inflow of Venezuelan migrants into Colombia through the land entry ports after 2016 and their location in municipalities closer to the border. I build on the idea that right-wing political parties could have profited from their presence during the campaign for the presidential election.

The ideal experiment would compare electoral outcomes in the 2018 election across municipalities of Colombia with as good as random variation in the inflow of Venezuelan migrants after the ports re-opened. However, new arriving migrants tend to settle in enclaves established by earlier immigrants, making their presence correlated with unobserved characteristics (Card, 2001; Jaeger, Ruist and Stuhler, 2018; Rozo and Vargas, 2021). Therefore, a comparison between municipalities that received higher and lower migration is likely to be biased.

I take this idea to the data by seeking exogenous factors that may have triggered such random variation. I construct an instrumental variable using the distance of each municipality to the closes land entry port to instrument for the presence of Venezuelan migrants. In particular, I calculate the minimum distance between the municipality and the five principal entry ports located in the shared border between the two countries. After 2016, most Venezuelans entered through one of this five border crossings points. Their presence increased throughout the territory covering about 95% of the municipalities but sorted more disproportionately in municipalities closer to the border (see Figure 3.3).

My identification strategy relies on the assumption that the distance of a municipality to the closest entry port predicts the presence of migrants, and it is not correlated directly with the electoral results of the 2018 elections and the political preferences in said municipality. The distance of a municipality to the closest entry port seems to explain the migration pattern of Venezuelan after 2016. Using information from the last two population censuses, I find that the municipalities closest to the border experienced the largest positive change in presence of Venezuelan migrants (see Figure 3.3). At the same time, distance to the port is unlikely to have a direct effect on how Colombians voted in the 2018 election. In my setting, I argue that any effect will take place indirectly through the presence of Venezuelan migrants. I formally test these assumptions in section 3.5.2.

There are some threats to my identification: (i) the distance of a municipality to the border is correlated with the political preferences in said municipalities, (ii) the preferences for political parties or ideologies can be geographically grouped, (iii) Venezuelan migrants could have chosen more welcoming municipalities as their destination or (iv) decided to migrate to enclaves established by early immigrants.

These concerns are partly alleviated by the political context in Colombia. Local support for political parties is based on a clientelist system of political exchange that emerged in the 1960's where the two main parties shared power and excluded other political movements (Fergusson et al., 2021). This bipartisan dominance collapsed with the adoption of the 1991 constitution—later modified by the 2003 political reform—that lowered entry barriers for new parties, which consolidated after the 2006 elections. Supporters of these new parties blurred any previous geographic concentration of preferences. Nonetheless, to account for potential differences in political preferences, I control for the electoral results of the previous presidential elections and for geographic characteristics. Furthermore, I control for the presence of migrants from other countries as a proxy for how welcoming the municipality is to migrants. Finally, in section 3.7, I show some analysis using early settlements to show whether the preferences of the inhabitants of a given municipality are different due to the long-term exposure to migrants.

3.5 Empirical Strategy

3.5.1 Empirical specification

The relationship of interest is the following:

$$Votes_m^{2018} = \alpha + \beta Ven_m + \sum_J \delta^j votes_m^j + \delta_g geo_m + \delta_X X_m + \varepsilon_m \tag{3.1}$$

where $Votes_m^{2018}$ is the share of votes for each one of the ideologies and each one of the candidates in the 2018 presidential election; Ven_m is the migration measure that counts the number of Venezuelan migrants that enter Colombia between 2013 and 2017; $votes_m^j$ are the share of votes for the right- and for the left-wing parties, separately, in election years $J = \{2002, 2006, 2010, 2014\}$; geo_m are geographic controls that include the distance of each municipality to the country capital (Bogotá), the Department's capital, and to the closest principal market, and five dummy variables that take the value of 1 if the municipality share a border with either of the 5 neighboring countries (Brazil, Ecuador, Panama, Peru, or Venezuela); and X_m are municipality characteristics including the area and elevation of the municipality, the population, the share of rural population as proxy of rurality, and the share of migrants that are not from Venezuela—to capture any unobserved heterogeneity related more amenable locations to migrants—using information from the 2005 census. ε_m is an idiosyncratic error term, β is the parameter of interest that captures the relation between the presence of Venezuelans and the preference for each one of the candidates.

I exploit the distance of each municipality to the entry port on the shared border to instrument for the presence of Venezuelan migrants. I take this idea to the data using a Two-Stage Least Squares (2SLS) with the following first-stage regression specification:

$$Ven_m = \psi + \gamma f(distance_m) + \sum_J \theta^j votes_m^j + \theta_g geo_m + \theta_X X_m + \eta_m \quad (3.2)$$

where $f(distance_m)$ is a spatial decay function of the distance. In particular, I use a logarithmic function of the distance of the municipality to the closest entry port on the shared border between the two countries. To ease interpretation, I standardized the measure to have zero mean and unit variance. The other variables were defined above and η_m is an idiosyncratic error term. The parameter γ captures the relevance of the instrument which should be positively correlated with the presence of migrants. I present evidence in the next section that my estimates are not driven by the functional form of the decay function.

3.5.2 Validation of the research design

To assess the validity of the instrument, the distance from the municipality to the closest entry port needs to explain the presence of Venezuelan migrants. The parameter γ in equation 3.2 captures this relation. Table 3.1 in Column (1) reports the estimates of γ by estimating equation 3.2 without including the control variables, i.e., imposing $\theta_j = \theta_g = \theta_X = 0$; while in Column (2), I relax the restriction by including the control variables. Columns (3) to (10) report the estimates using different decay functions. Odd columns present estimates that do not include any control variables and even columns those that include the full set of controls. The last four columns use the distance of each municipality to the main entry port located at Cúcuta, the Simón Bolívar International Bridge, which could suggest that the entry port is not related to any unobserved characteristic of the municipalities in Colombia instead of to the closest entry port.

Estimates present a positive and significant coefficient for the inverse logarithmic and the inverse linear function, as expected. For a linear decay function, the estimates present a negative and significant relation.[12] In all the cases, the estimated coefficient is around 0.3 percentage points in absolute value. For instance, the coefficient from Column (2) implies that a one standard deviation increase in distance measure increases the presence of Venezuelan migrants by 0.4 percentage points. The F-test for excluded instruments shows the relevance of the instrument for the different decay functions, except for the inverse linear case.

To further assess the validity of the exclusion restriction, I investigate whether there are underlying political preferences. associated with the distance between the municipalities and the Venezuelan frontier. I divide my analysis in two parts: first, I use previous election as placebos to test whether my measure of Venezuelan migrants is correlated with political preferences in previous elections. I estimate equation 3.1 using the share of votes in previous elections as the outcome variable. Note, given that I am using $votes^j$ as outcome variables, I impose the restriction that $\theta_j = 0$ for all $j \in J$ but include the other control variables. Second, I investigate whether distance is correlated with electoral outcomes in previous elections.

Figure 3.4 report the OLS estimates in panel a. They suggest that the presence of Venezuelan migrants is correlated with political preferences. However, notice that in 2006 the right-wing candidate, Alvaro Uribe, was running for re-election and he was elected with more than 60% of the votes. Given his large popularity, it is possible that there is a spurious correlation between the future presence of migrants and his share of votes. Similarly, the estimates for the share of votes for Juan Manuel Santos suffer from a similar problem as he was the appointed successor of Uribe. Santos had around 46% of the total votes and double the second highest voted candidate, Mockus with 21% of the total votes. On the other hand, the reduce form estimates reported in panel b

[12] I show that my estimates are robust to the use of an exponential decay function in Appendix B.1, Figure B.1.

suggest that the distance from each municipality to the closest entry port does not directly explain the political preferences in those municipalities as estimates are not stable from one election to the next.

Taken together, these results suggest that the presence of Venezuelan migrants captures some historical political preferences, but they are not driven by my measure. Therefore, my research design relies on controlling for the results of the previous election for identification, as there could be some unobserved political preferences that could affect my main estimates. Intuitively, I will be comparing municipalities closer to the border with those further away that have similar political preferences.

3.6 Results

3.6.1 Main results

My main results are summarized in Figure 3.5. It shows the estimates of the presence of Venezuelan migrant on the share of votes for every candidate and for every political ideology. Tables 3.2 and 3.3 report these coefficients. Panel A presents the estimates of equation 3.1 using a 2SLS with the distance of each municipality to the closest migration port as the instrument for the presence of Venezuelan migrants. Panel B present the estimates of equation 3.1 using OLS, while Panel C present the reduce form results of the instrument on the outcome variable. Panel D presents the first stage estimates from equation 3.2. Odd columns estimates do not include any control variable and even columns include the full set of controls. My preferred specification is reported in Panel A and includes the full set of controls. I present the results for the winning and right-wing candidate first, Duque, then I complement my analysis including the other candidates and the political ideologies.

My empirical findings suggest that a one percent increase in the presence of Venezuelan migrants increase the share of votes for the leading right-wing party, *Centro Democrático*, by 8 percentage points (see Column (4) in Table 3.2). This is a sizable effect as the average share of votes for the party at the municipal level is 46.9%. It is important to highlight that these results must be taken with caution as the estimates are not very precise due to the low variance of the explanatory variable (standard deviation is less than 1%). This lower variance leads to point estimates that are somewhat large. For instance, if I were to take a municipality without presence of Venezuelan migrants and input the maximum value, which is about 7%, then the share of votes for Duque would have increase by 56 percentage points, an unlikely scenario. I explore whether this result is driven by the measure of presence of Venezuelan migrants in section 3.7.

The results of estimating equation 3.1 using OLS and the full set of controls (Column (4) in Panel B) suggests that an increase of 1 percent in the presence of migrants increase the share of votes for Duque in 1.5 percentage points. Note, IV estimates are larger than those of OLS, i.e., OLS estimates are biased toward zero. The difference in point estimates between these two estimates comes from two potential sources. First, recent migrants may have located in early Venezuelan settlements and the presence of these early migrants could have affected the long-term political preferences of the inhabitants of the recipient municipality (Jaeger, Ruist and Stuhler, 2018; Rozo and Vargas, 2021). Second, the measure of presence of Venezuelan suffer from measurement error, as explained above. However, both of these sources of bias are solved through the use of instrumental variables as long as the instrument—i.e., the distance to the land entry port—is not correlated with neither the unobserved preferences nor the measurement error term. In section 3.7, I explore the difference by presence of early settlements.

I analyze how the presence of Venezuelan migrants affected the votes for each candidate and I find that the presence of Venezuelan migrants polarized the election by increasing the votes for the Iván Duque at the expense of the votes of Gustavo Petro (see Figure 3.5). This resonate with the results at the ideology level by suggesting that the presence of Venezuelan migrants increased the share of votes for the right-wing parties at the expense of the share of votes for the left-wing candidates. This finding is similar to that of Rozo and Vargas (2021) who suggest a shift in the votes from the left- towards the right-oriented candidates. Differently from previous studies, my results also suggest that the increase in the share of votes for the right-wing parties come exclusively from the increase in the share of votes for Iván Duque, the *Centro Democrático* party candidate. This finding marks a difference with the previous literature that has associated the increasing share of votes for right-wing parties with higher presence of migrants (Barone et al., 2016; Dustmann, Vasiljeva and Damm, 2018; Edo et al., 2019; Gerdes and Wadensjö, 2010; Halla, Wagner and Zweimüller, 2017; Mayda, Peri and Steingress, 2022; Otto and Steinhardt, 2014).

The estimates on turnout explore whether the increase in the share of votes come from more citizens going to vote rather than a change in preferences of existing voters (see Column (10) in Table 3.2). In this case, the results suggest that the effect come via the two channels mentioned, as the point estimates is not as large as the one for the share of votes. First, the presence of Venezuelans had a positive effect on the participation in the election by increasing the number of voters and, second, by shifting the preferences of the voters away from the left-oriented candidates and towards the leading right-oriented candidate.

Estimates for the second round of the presidential election are consistent with those of the first-round. Notice, the run-off election was held between the two candidates with the highest votes in the first round, namely: Iván Duque and Gustavo Petro. These two candidates represent the polar opposites as their platforms have been associated with extreme right position and with left position (Gallego et al., 2019). My estimates for the effect of the presence of Venezuelan migrants in the second round resonate those of the first round and are presented in Table B.3 in Appendix B.1. The share of votes for the right-wing party candidate, Iván Duque, increases at the expense of the left-wing candidate, Gustavo Petro.

3.7 Robustness

3.7.1 Migration measure

One potential concern is that my results are driven by the measure of presence of Venezuelan migrants. My measure of presence of Venezuelan migrants counts the number of Venezuelans that arrive between the first semester of 2013 and the first semester of 2017. I use two alternative measures of presence and report the estimates in Table 3.4. The first two columns present the main estimates to ease comparison. The next two columns measure the presence of Venezuelan migrants as those who entered the country after the first semester of 2017 or within the last year; while the last two columns measure it as those who entered the country after the first semester of 2013 or within the last five years.

The estimates are robust to different migration measures. In particular, they are larger and less precise when using only those that arrive in the previous year, while the estimates using the larger measure is more precise and about two thirds of the main estimate. Notice, the migration measure using only the individuals that arrive within the last year has lower explanatory variable as it has a small range (min is zero and max is 5%) and low variance (SD=0.5). In contrast, the broader measure of presence used in the last two columns has double the range (same min and 10.6% max) and a larger variance (SD=1.3) giving it more explanatory power. My preferred measure, the one in the first two columns, is located somewhere in the middle. These features help to explain why the estimates are larger when using more constraint definitions. Nonetheless, it is reassuring that the estimates point in the same direction and do not lose significance, nor do the first-stage estimates change in the explanatory power.

3.7.2 Political preferences and Centro Democrático

To further validate my design, I explore whether my results are driven by a preference for the *Centro Democrático* party. Note, I cannot extend this analysis to the left-wing candidate, Gustavo Petro, as he run for a party created in 2011 that did not take part in the 2014 presidential election, therefore, I cannot capture preferences for his party. I calculate the change in votes for *Centro Democrático* party between presidential election and rule out the possibility of a preference for the party. Notice, similar to the part of Petro, I can only do this with the 2014 presidential election result as the party did not participate before in any election. In addition, I calculate the difference between the share of votes for Duque and the votes for the NO option in the 2016 peace plebiscite as the party was leading the opposition to the adoption of the peace agreement. Table 3.5 reports the IV estimates using these two alternative outcomes. Note, since I am using information from the 2014 election, I do not control the electoral results of that year (reported in Columns (3) and (4)). These results are remarkably similar to those of the main specification, suggesting that there is no unobserved characteristic related to ideology that is driving the results. However, the estimates for the change with respect to the peace agreement plebiscite result are smaller, which suggests that the attitudes towards the peace agreement help explain the electoral results of 2018. I explore this result further in section 3.9.

3.7.3 Early settlements

It could be the case that municipalities that have been exposed to Venezuelan migrants before reacting differently from those where migrants are arriving for the first time. To analyze this, I use information on the location of Venezuelan migrants using the 2005 Colombian census. To keep simple the analysis, I split the sample in two groups: those that had a prior presence of Venezuelan migrants in 2005 and those who had not; interestingly, each subsample accounts for roughly half of the municipalities. Table 3.6 reports the results for these two samples and the totality of the municipalities , to ease comparison. The IV estimates in the municipalities with previous settlements presents a similar point estimate that for the full sample validating my results. However, the IV estimates for the sample without prior presence is about 3.5 times larger (Column (6)) when compared to those of the full sample. These results suggest that voters that were not exposed in the past reacted more to the presence of the Venezuelan migrants than those that had been exposed in the past.

3.8 Additional Results

3.8.1 Labor market

I follow the literature and explore the relationship between presence of migrants and labor market outcomes in Colombia (Bahar, Ibáñez and Rozo, 2021; Lebow, 2022; Santamaria, 2020). These studies exploit the large migration wave and work permissions granted to Venezuelan.

Their findings suggests that there are negligible effects on labor market outcomes. I use the question on the census that asked about last week's activity to approximate the labor market. Figure 3.7 summarizes the results for the inhabitants of the municipality and only for the Venezuelans. The results present estimates not statistically different from zero, although the share of individuals working decreased and that of studying increased slightly. However, the results must be taken with caution as not all the individuals responded to the questionnaire during the same time, making it harder to disentangle the effects from temporal variation or economic cycles.

3.8.2 Education level

Recent literature has studied the role of education and skill level on the heterogeneous effects of immigration (Edo et al., 2019; Mayda, Peri and Steingress, 2022). Their results suggest that the support for right-wing candidates tend to be favored by the presence of low educated (skilled) migrants. To explore this channel, I use the questions on the census that ask about gender, whether the respondent knows how to read and write, and the highest education level attained. Figure 3.8 summarizes the results for the inhabitants of the municipality and for the Venezuelans. The results suggest that there is an increase in the share of respondents that have completed college, and a decrease in the share of individuals that have completed either middle- or high-school. The estimates for the Venezuelan respondents suggest that this effect does not come from their presence in the municipality. A possible explanation is that migrants decide to locate in the biggest cities and more urban (less rural) areas where there is more demand for college educated individuals.

Furthermore, the estimates for Venezuelan respondents only suggest that there is a significant increase in the number of Venezuelans that have completed only primary school and a negative, although not significant, for those in higher education levels. This result goes in line with a recent characterization of Venezuelan migrants based on the Colombian households' surveys of 2015 and 2016. These surveys suggests that 80 percent of registered migrants have not completed a high school education, and at least half of them are 25 years old or less (OLR, 2017). Moreover, according to Rozo and Vargas (2021) and Knight and Tribin (2023), the recent migrant wave was composed by poor, vulnerable, and poorly educated households.

3.9 Mechanisms

3.9.1 Violent crime

I rule out that the change in voting preferences is due to changes in violent crimes. Figure 3.6a and Figure 3.6b presents the estimates using person robbery and homicide rates, respectively, from 2011 to 2017. To account for unobserved characteristics, I compute the outcome as the deviation between the robbery (homicide) rate for every year with respect to that of 2010. Each one of the estimates comes from a different IV regression using these outcomes. To account for any underlying relation that my measure could be capturing, I normalize the values for each year by calculating the deviation with respect to the information from 2010. In both cases, the person robbery and the homicides present some fluctuation but are not statistically distinguishable from zero in most of the years. These results suggest that there is no effect on violent crimes, contrary to the findings of Knight and Tribin (2023) who find an increase in homicides in Colombia after the reopening of the border in 2016.

3.9.2 Mayoral elections

The presence of migrants could have affected political preferences at the local level, which in turn could have an effect in the presidential election. To analyze this, I explore whether the presence of Venezuelan migrants affected the political preferences in the 2015 local elections. Table 3.7 presents the effects of the Venezuelan migrants on the votes for the right-, left-, and center-oriented candidates. I estimate equation 3.1 using the share of votes for each of the factions as outcome. It is not possible to do this analysis at the party level as at the local level there are many independent candidates (Woldemikael, 2022). I find a positive effect on the share of votes for right-wing candidates at the expense of the votes for the center-oriented candidates. This result contradict the findings of Rozo and Vargas (2021) who find that the center-oriented candidates are not affected by the presence of migrants whereas left-wing candidates are.

3.9.3 Peace Plebiscite

The 2018 presidential election was marked by the adoption of the peace agreement. I study whether my results are driven by the preferences over this topic as the *Centro Democrático* party was the major political force opposing its adoption (Gallego et al., 2019). Furthermore, *Centro Democrático* use the example of Venezuela as a failed state during these two electoral processes. Figure 3.1 provide some evidence of the increasing interest of the term "castrochavista" around the dates of the processes. The last two columns of Table 3.7 present the estimates. There is a positive effect of the presence of migrants on the share of votes for the No option. Note, this is consistent with the support for the *Centro Democrático* party and because the measure of presence of migrants includes migrants that arrived before the plebiscite took place.

3.9.4 Mediation Analysis

I find that the presence of Venezuelan migrants have an effect on the votes for the local election and the peace plebiscite but not with the violent crimes. To understand how the preferences for presidential candidates are affected by these mediating factors, I introduce them in the model specified in equation 3.1 in order to create a "full" model (equation 3.3). Note, these variables could be seen as 'bad controls' as they are also affected by the presence of migrants, their relationship with the voting preference in the presidential election is not clear. Equation 3.1 is modified as follows:

$$Votes_m^{2018} = \alpha_M + \beta_M Ven_m + \sum_K \mu^k p_m^k + \sum_J \delta_M^j votes_m^j + \delta_{Mg} geo_m + \delta_{MX} X_m + \varepsilon_m \quad (3.3)$$

where I denote each of the mechanisms as p, i.e., $K = \{$the violent crimes, the votes for the right-oriented candidates in the local election, center-oriented candidates in the local election, votes for the No option in the peace plebiscite$\}$. I omitted the share of votes for the left-oriented candidates as it is collinear with the other shares of votes, and it present a non-significant relation.

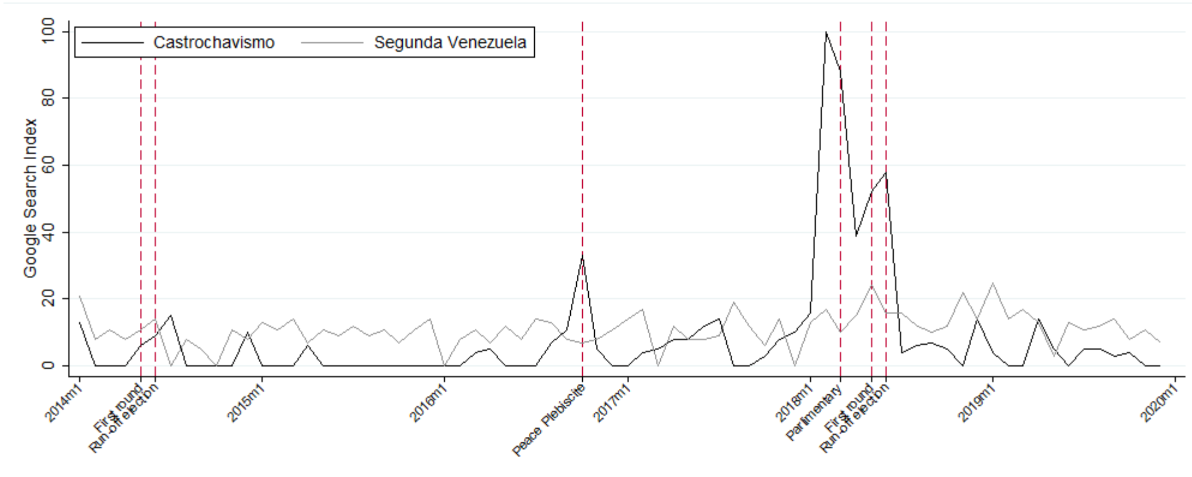
To quantify the importance of these mechanisms, Gelbach (2016) provides a decomposition to explain the relative importance of each one of these. Table 3.8 present the estimates for: violent crimes, votes for the right-oriented candidates in the local election, center-oriented candidates in the local election, votes for the No option in the peace plebiscite. To simplify the analysis, I use as outcome the share of votes for Iván Duque, *Centro Democrático* candidate. The results suggest that the preferences over the peace agreement explain all the difference between the "full" model and the base model from equation 3.1, and about a third of the total effect of the presence of Venezuelan migrants on the votes for the leading right-wing party candidate, Duque. This result provide further evidence that the attitudes towards the peace agreement were determinants of the electoral results of the 2018 presidential election.

3.10 Conclusion

This paper contributes to the literature of migration and, in particular, to the growing literature of the effects of migration inflow of Venezuelans. The case of Colombia allows me to identify the effects of the migration shock of Venezuelan on voting behavior for each candidate. My empirical strategy approximates the shock using data from the last national census and exploits the likeliness of new migrants settling in enclaves established by earlier migrants using data from the previous to last available census.

I find robust evidence suggesting that there is a positive relation between the share of Venezuelan migrants with the voting preference for a right-wing party candidate, Iván Duque from *Centro Democrático*. Contrary to the findings of the literature, I find that not all right-wing parties benefit from such presence. I also find that the presence of Venezuelan migrants polarized the election by increasing the share of votes for Centro Democrático at the expenses of the votes for the left-wing candidate, Gustavo Petro. These results support the strategic electoral misinformation hypothesis (Rozo and Vargas, 2021) where political parties demonize electoral competitors by associating them with political regimes.

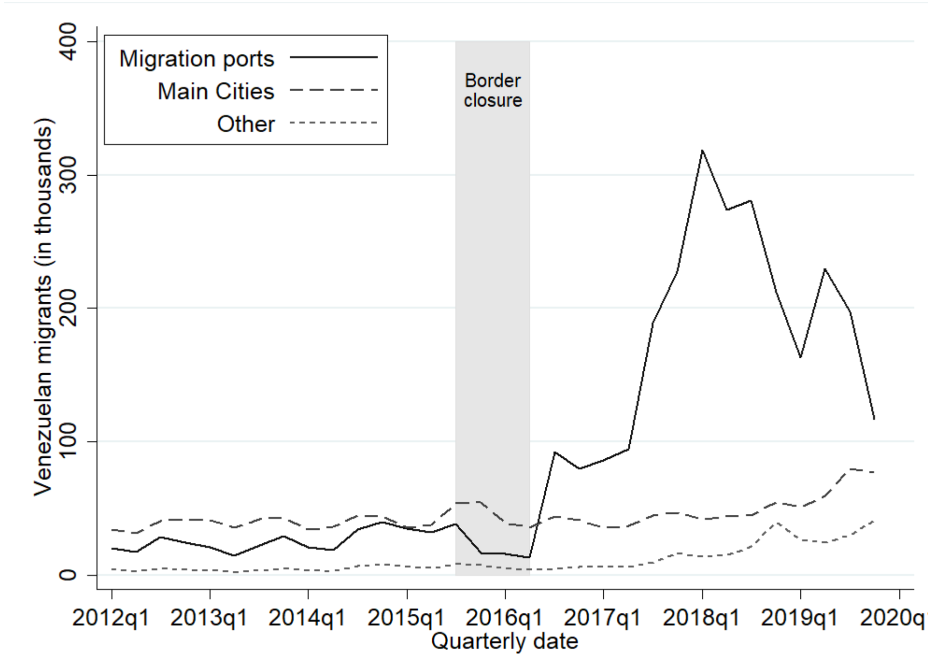
Figure 3.1: Google search for 'Castrochavismo' and 'Segunda Venezuela' in Colombia



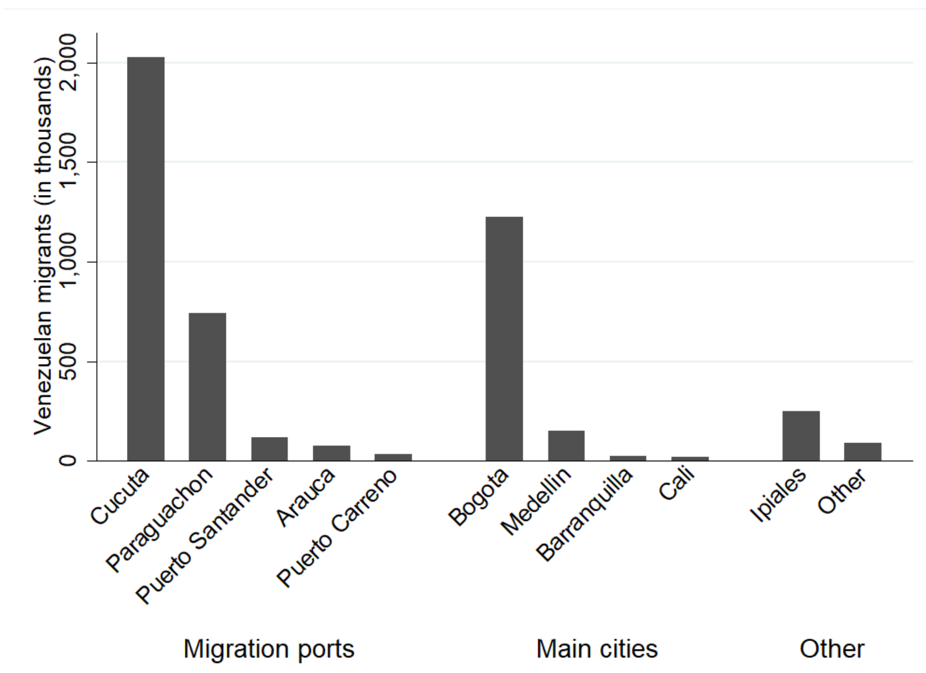
Source: Google trends. Chart present the number of Google hits for the search terms 'castrochavismo' and 'segunda Venezuela' in Colombia between for the years 2016 to 2020. Data is monthly scaled to range from 0 to 100, where 100 is the highest possible amount of searches.

Figure 3.2: Venezuelan migration to Colombia

(a) Over time

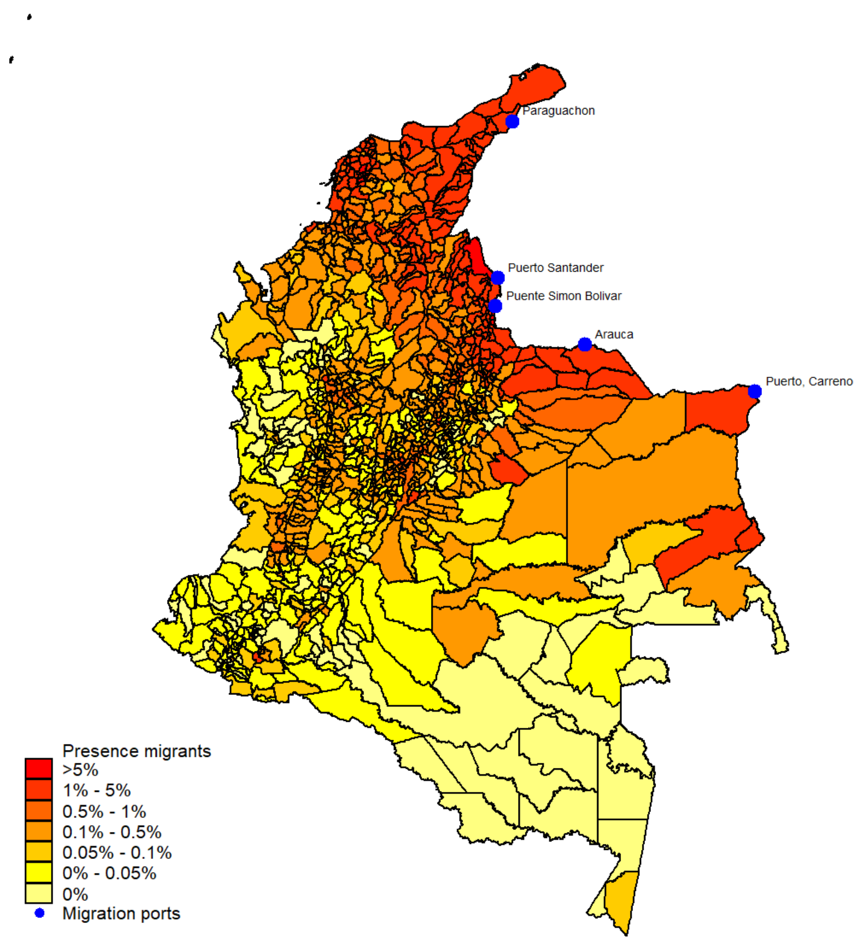


(b) By entry ports

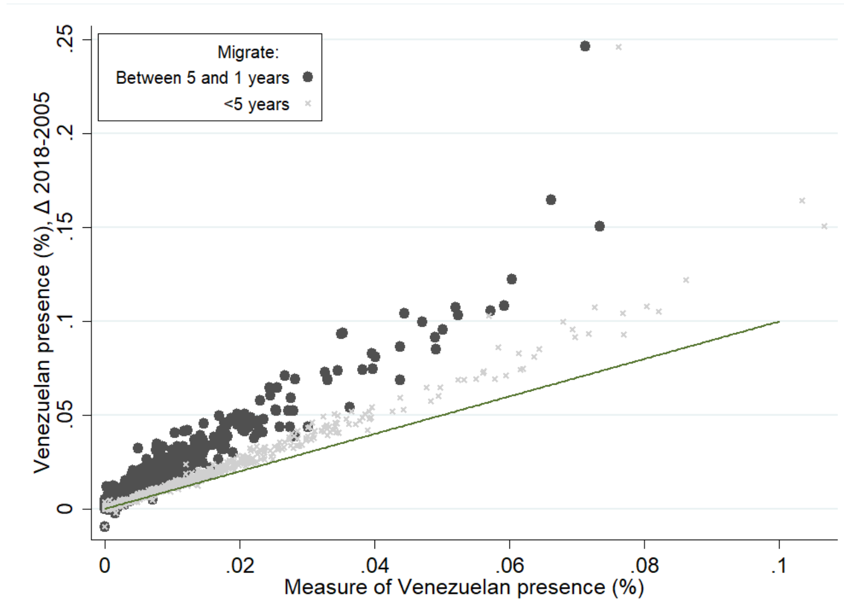


Source: migration office. Lines in panel a represent the number of Venezuelans entering Colombia through different entry ports from the first quarter in 2012 to the second quarter in 2019. The shaded area represents the border closure between August 2015 and August 2016. Panel b presents aggregates by entry port from the first quarter in 2012 to the second quarter in 2019.

Figure 3.3: Location of Venezuelan migrants in Colombia



(a) Measure of presence

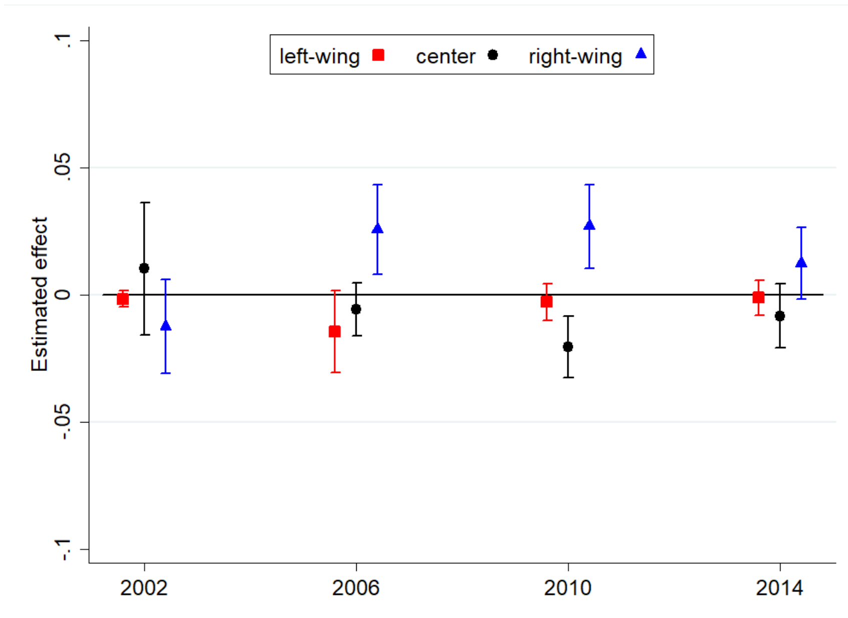


(b) Presence variation 2018-2005

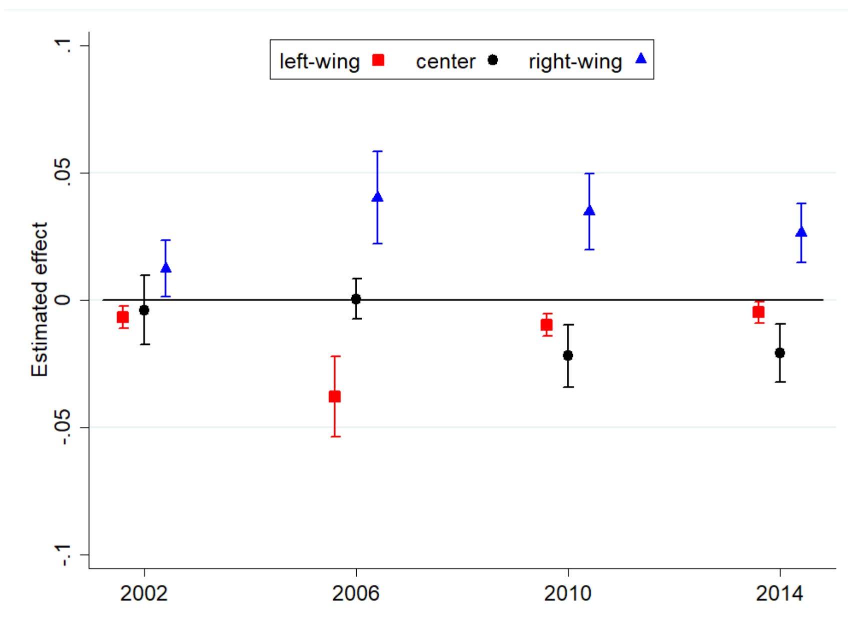
Source: Censuses from the Statistics Department. Panel a presents the measure of presence of Venezuelan migrants in the main analysis. Dots represent the location of the land migration ports. Panel b presents variation in the presence of Venezuelan migrants between 2018 and 2005 using censuses information in the vertical axis. The horizontal axis present the measure of presence of Venezuelan migrants in the main analysis. Solid line is a 45-degree line.

Figure 3.4: Presence of Venezuelan migrants and distance on previous presidential elections

(a) OLS estimates

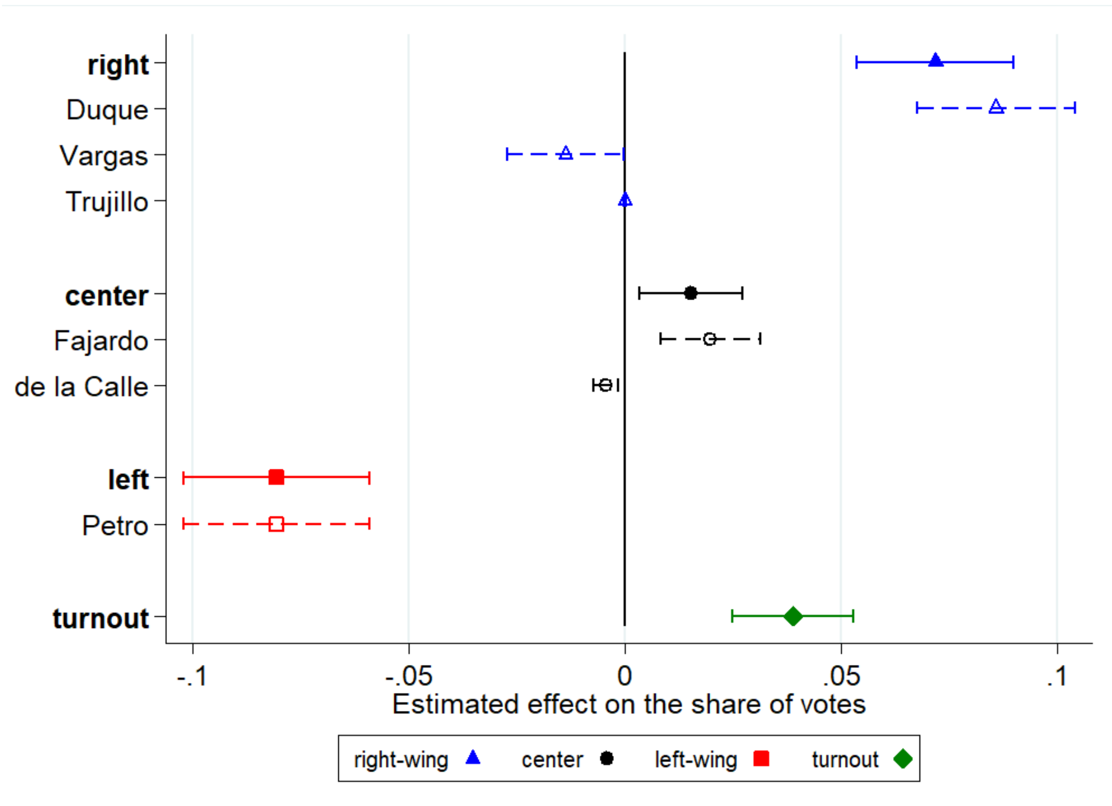


(b) Reduce form



Note: Panel 3.4a presents the OLS estimate of β in equation 3.1 using the votes for each ideology in the election year reported in the horizontal axis. Panel 3.4b presents the reduce form estimates using distance from each municipality to the closest entry port as the explanatory variable. Each point comes from a different regression.

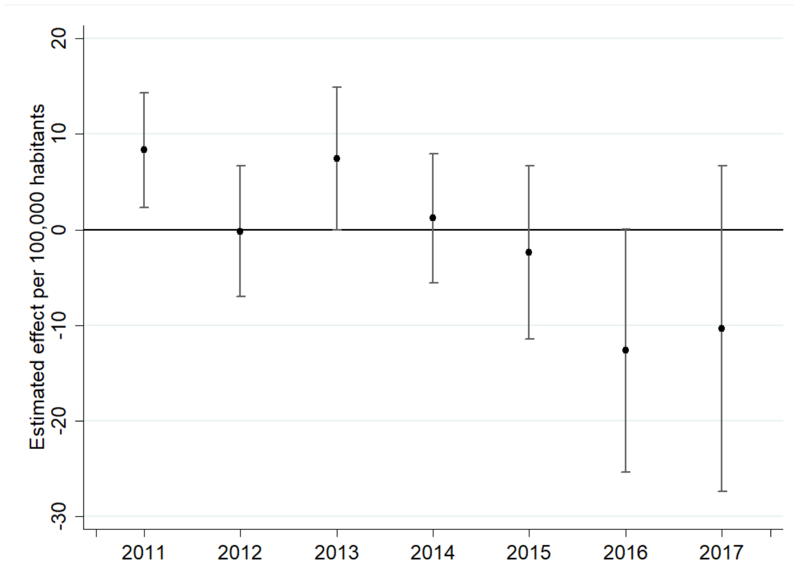
Figure 3.5: Effects of presence of Venezuelans on the votes of the 2018 presidential election



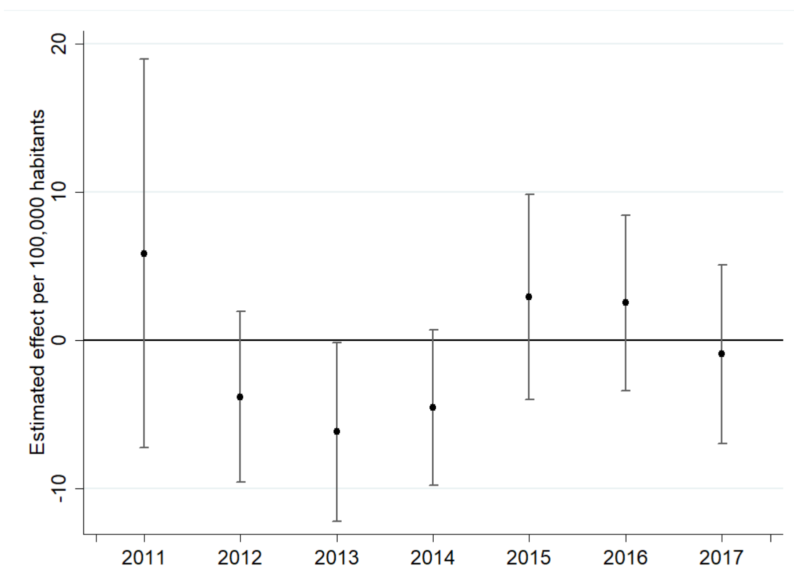
Note: Each point represents the IV estimate of β of equation 3.1. Estimates in solid symbols represent the aggregated value for each political ideology and for the turnout rate. Estimates in hollow symbols represents individual candidates.

Figure 3.6: Effects of presence of Venezuelan migrants on violent crime

(a) Person robbery

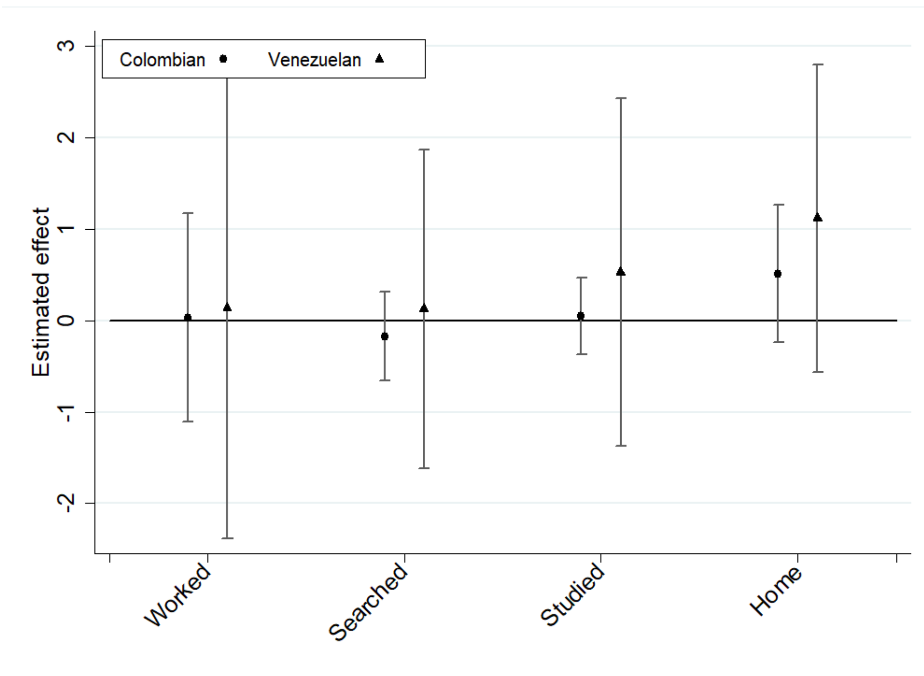


(b) Homicide



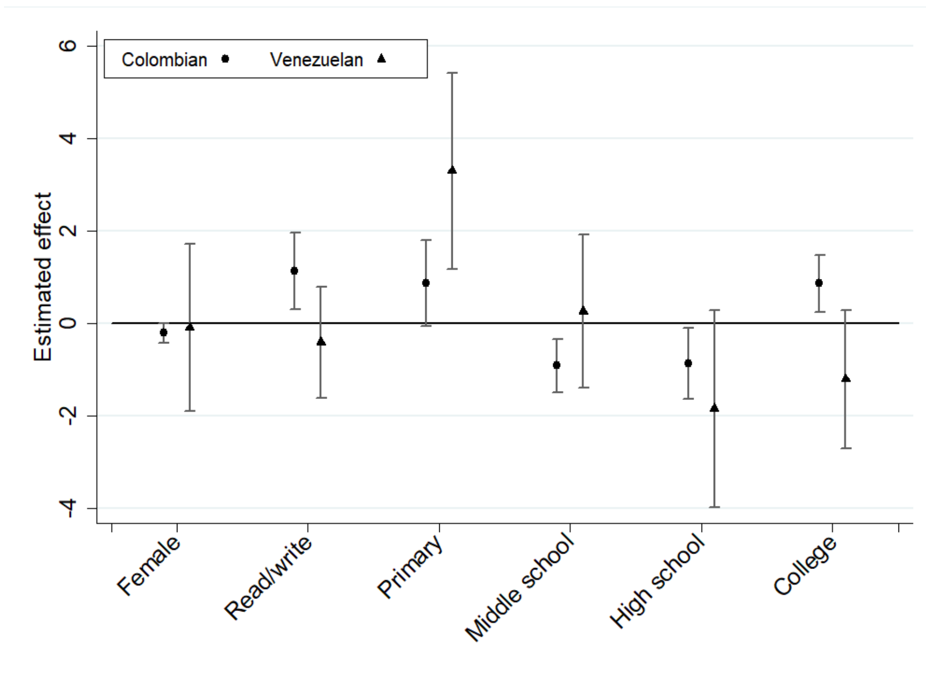
Note: Each point represents the IV estimate of β of equation 3.1 using violent crimes as outcomes. Panel 3.6a present the estimates for person robbery, while Panel 3.6b for homicides. Both rates are calculated as deviation from its value in 2010.

Figure 3.7: Effects of Venezuelan presence on labor market



Note: Each point represents the IV estimate of β of equation 3.1 using labor outcomes as the dependent variable. Labor market outcomes are calculated using census data for Colombian and Venezuelan separately. Circles represent estimates using outcomes for the Colombian population only and triangles for Venezuelan population only.

Figure 3.8: Effects of Venezuelan presence on education level



Note: Each point represents the IV estimated coefficient of equation 3.1 using educational characteristics as the dependent variable. Educational characteristics are calculated using census data for Colombian and Venezuelan separately. Circles represent estimates using outcomes for the Colombian population only and triangles for Venezuelan population only.

Table 3.1: First-stage: Presence of Venezuelan migrants as a function of distance

Distance function:	Minimum distance to 5 entry ports						Distance to Cucuta			
	Inverse Logarithmic		Inversee Linear		Linear		Inverse Logarithmic		Linear	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Distance function	0.005***	0.004***	0.004**	0.002**	-0.004***	-0.003***	0.004***	0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
F-test (excluded instruments)	69.880	41.078	6.545	4.649	190.197	141.939	122.850	105.264	150.085	151.617
Observations	1,122.000	1,094.000	1,122.000	1,094.000	1,122.000	1,094.000	1,122.000	1,094.000	1,122.000	1,094.000
Border controls		✓		✓		✓		✓		✓
Distance controls		✓		✓		✓		✓		✓
Administrative controls		✓		✓		✓		✓		✓
Electoral controls		✓		✓		✓		✓		✓

Notes: Each column represents the estimate of γ in equation 3.2. Dependent variable is the presence of Venezuelan migrants. Independent variables is a decay function of the distance of each municipality to an entry port. Functional form and entry port are reported in the column header. Odd Columns do not include any control and even columns include all the control variables. Robust standard errors presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 3.2: Effects of presence of Venezuelans on the votes of the 2018 presidential election

Dependent variable:	Votes for right		Right candidates						Total votes	
			Votes for Duque		Votes for Vargas		Votes for Trujillo			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: 2SLS IV regression (second-stage)										
Presence Venezuelan migrants	0.105***	0.063***	0.102***	0.079***	0.003	-0.016**	-0.000	-0.000	0.019**	0.040***
	(0.013)	(0.008)	(0.014)	(0.013)	(0.004)	(0.008)	(0.000)	(0.000)	(0.008)	(0.014)
Panel B: OLS regression										
Presence Venezuelan migrants	0.026***	0.023***	0.007	0.015***	0.019***	0.009*	0.000	-0.000***	-0.000	0.021***
	(0.007)	(0.004)	(0.007)	(0.005)	(0.005)	(0.005)	(0.000)	(0.000)	(0.003)	(0.005)
Panel C: Reduced form										
Distance function	0.056***	0.024***	0.055***	0.030***	0.001	-0.006*	-0.000	-0.000	0.010**	0.015**
	(0.010)	(0.005)	(0.011)	(0.008)	(0.002)	(0.004)	(0.000)	(0.000)	(0.005)	(0.007)
Panel D: 2SLS first-stage regression										
Distance function	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F-test (excluded instruments)	69.88	41.08	69.88	41.08	69.88	41.08	69.88	41.08	69.88	41.08
Observations	1,122	1,094	1,122	1,094	1,122	1,094	1,122	1,094	1,122	1,094
Border controls		✓		✓		✓		✓		✓
Distance controls		✓		✓		✓		✓		✓
Administrative controls		✓		✓		✓		✓		✓
Electoral controls		✓		✓		✓		✓		✓

Notes: Each entry is estimated separately. Panel A and B present estimates of equation 3.1 and panel D of equation 3.2. Dependent variables are reported in the column header. First-stage estimates uses an inverse logarithmic decay function of the distance to the closest entry port. Odd Columns do not include any control and even columns include all the control variables. Robust standard errors presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 3.3: Effects of presence of Venezuelans on the votes of the 2018 presidential election (continued)

Dependent variable:	Votes for center		Center candidates: Votes for Fajardo		Center candidates: Votes for de la Calle		Votes for left		Left candidates: Votes for Petro	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: 2SLS IV regression (second-stage)										
Presence Venezuelan migrants	-0.011***	0.006	-0.006	0.012	-0.005***	-0.005***	-0.088***	-0.062***	-0.088***	-0.062***
	(0.004)	(0.007)	(0.004)	(0.008)	(0.001)	(0.001)	(0.013)	(0.012)	(0.013)	(0.012)
Panel B: OLS regression										
Presence Venezuelan migrants	-0.016***	0.004	-0.015***	0.004*	-0.001	-0.001	-0.006	-0.024***	-0.006	-0.024***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.006)	(0.004)	(0.006)	(0.004)
Panel C: Reduced form										
Distance function	-0.006***	0.002	-0.003*	0.004	-0.003***	-0.002***	-0.047***	-0.024***	-0.047***	-0.024***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.010)	(0.008)	(0.010)	(0.008)
Panel D: 2SLS first-stage regression										
Distance function	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F-test (excluded instruments)	69.88	41.08	69.88	41.08	69.88	41.08	69.88	41.08	69.88	41.08
Observations	1,122	1,094	1,122	1,094	1,122	1,094	1,122	1,094	1,122	1,094
Border controls		✓		✓		✓		✓		✓
Distance controls		✓		✓		✓		✓		✓
Administrative controls		✓		✓		✓		✓		✓
Electoral controls		✓		✓		✓		✓		✓

Notes: Each entry is estimated separately. Panel A and B present estimates of equation 3.1 and panel D of equation 3.2. Dependent variables are reported in the column header. First-stage estimates uses an inverse logarithmic decay function of the distance to the closest entry port. Odd Columns do not include any control and even columns include all the control variables. Robust standard errors presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 3.4: Effects of Venezuelan presence using different migration measures

Dependent variable: Votes for Duque	Migration measure		Migrate <1 year		Migrate <5 years	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2SLS IV regression (second-stage)						
Presence Venezuelan migrants	0.102***	0.079***	0.209***	0.175***	0.069***	0.055***
	(0.014)	(0.013)	(0.028)	(0.031)	(0.008)	(0.006)
Panel B: OLS regression						
Presence Venezuelan migrants	0.007	0.015***	0.005	0.017***	0.004	0.010***
	(0.007)	(0.005)	(0.011)	(0.006)	(0.005)	(0.003)
Panel C: Reduced form						
Distance function	0.055***	0.030***	0.055***	0.030***	0.055***	0.030***
	(0.011)	(0.008)	(0.011)	(0.008)	(0.011)	(0.008)
Panel D: 2SLS first-stage regression						
Distance function	0.005***	0.004***	0.003***	0.002**	0.008***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F-test (excluded instruments)	69.88	41.08	17.81	6.55	42.69	20.16
Observations	1,122	1,094	1,122	1,094	1,122	1,094
Border controls		✓		✓		✓
Distance controls		✓		✓		✓
Administrative controls		✓		✓		✓
Electoral controls		✓		✓		✓

Notes: Each entry is estimated separately. Panel A and B present estimates of equation 3.1 and panel D of equation 3.2. Dependent variable is the share of votes for Iván Duque. Independent variable is reported in the column header. First-stage estimates uses an inverse logarithmic decay function of the distance to the closest entry port. Odd Columns do not include any control and even columns include all the control variables. Robust standard errors presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 3.5: Effects of Venezuelan presence on votes for Centro Democratico

Dependent variable:	Votes for CD/Duque		Δ CD = Duque - Zuluaga		Δ CD = Duque - NO	
	(1)	(2)	(3)	(4)	(5)	(6)
		Panel A: 2SLS IV regression (second-stage)				
Presence Venezuelan migrants	0.102***	0.079***	0.064***	0.075***	0.038***	0.014*
	(0.014)	(0.013)	(0.009)	(0.013)	(0.009)	(0.008)
		Panel B: OLS regression				
Presence Venezuelan migrants	0.007	0.015***	0.018***	0.016***	0.010***	0.005
	(0.007)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
		Panel C: Reduced form				
Distance function	0.055***	0.030***	0.035***	0.029***	0.021***	0.005
	(0.011)	(0.008)	(0.007)	(0.008)	(0.006)	(0.004)
		Panel D: 2SLS first-stage regression				
Distance function	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F-test (excluded instruments)	69.88	41.08	69.88	41.43	69.88	41.08
Observations	1,122	1,094	1,122	1,094	1,122	1,094
Border controls		✓		✓		✓
Distance controls		✓		✓		✓
Administrative controls		✓		✓		✓
Electoral controls		✓		✓		✓

Notes: Each entry is estimated separately. Panel A and B present estimates of equation 3.1 and panel D of equation 3.2. Dependent variables are reported in the column header. First-stage estimates uses an inverse logarithmic decay function of the distance to the closest entry port. Odd Columns do not include any control and even columns include all the control variables. Robust standard errors presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 3.6: Effects of Venezuelan presence by presence of early settlements

Dependent variable: Votes for Duque	Full sample		Previous migration		No previous migration	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2SLS IV regression (second-stage)						
Presence Venezuelan migrants	0.101***	0.079***	0.086***	0.078***	0.872***	0.280***
	(0.014)	(0.013)	(0.013)	(0.014)	(0.179)	(0.071)
Panel B: OLS regression						
Presence Venezuelan migrants	0.006	0.015***	0.004	0.014***	0.030	0.003
	(0.007)	(0.005)	(0.008)	(0.005)	(0.034)	(0.016)
Panel C: Reduced form						
Distance function	0.054***	0.030***	0.044***	0.027***	0.181***	0.061***
	(0.011)	(0.008)	(0.010)	(0.008)	(0.018)	(0.010)
Panel D: 2SLS first-stage regression						
Distance function	0.005***	0.004***	0.005***	0.003***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
F-test (excluded instruments)	69.72	41.08	45.99	25.65	44.39	28.44
Observations	1,113	1,094	555	551	558	543
Border controls		✓		✓		✓
Distance controls		✓		✓		✓
Administrative controls		✓		✓		✓
Electoral controls		✓		✓		✓

Notes: Each entry is estimated separately. Panel A and B present estimates of equation 3.1 and panel D of equation 3.2. Dependent variable is the share of votes for Iván Duque. First-stage estimates uses an inverse logarithmic decay function of the distance to the closest entry port. Odd Columns do not include any control and even columns include all the control variables. Subsample reported in column header. Robust standard errors presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 3.7: Effects of presence of Venezuelan migrants on previous electoral cycles

	2015 mayoral election								Plebiscite	
Dependent variable:	Votes for right		Votes for center		Votes for left		Total votes		NO option	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: 2SLS IV regression (second-stage)									
Presence Venezuelan migrants	0.038***	0.050***	-0.007	-0.054**	-0.024*	0.002	0.070***	0.039**	0.064***	0.066***
	(0.013)	(0.018)	(0.015)	(0.021)	(0.012)	(0.015)	(0.016)	(0.016)	(0.009)	(0.011)
	Panel B: OLS regression									
Presence Venezuelan migrants	0.008	0.018*	-0.008	-0.030**	-0.001	0.003	0.027***	0.013	-0.003	0.010*
	(0.008)	(0.011)	(0.010)	(0.013)	(0.008)	(0.009)	(0.008)	(0.010)	(0.007)	(0.006)
	Panel C: Reduced form									
Distance function	0.023***	0.023***	-0.004	-0.025**	-0.014**	0.001	0.042***	0.018**	0.035***	0.025***
	(0.007)	(0.008)	(0.009)	(0.010)	(0.007)	(0.007)	(0.009)	(0.007)	(0.006)	(0.006)
	Panel D: 2SLS first-stage regression									
Distance function	0.006***	0.005***	0.006***	0.005***	0.006***	0.005***	0.006***	0.005***	0.005***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
F-test (excluded instruments)	163.91	126.35	163.91	126.35	163.91	126.35	163.91	126.35	69.88	41.08
Observations	1,035	1,020	1,035	1,020	1,035	1,020	1,035	1,020	1,122	1,094
Border controls		✓		✓		✓		✓		✓
Distance controls		✓		✓		✓		✓		✓
Administrative controls		✓		✓		✓		✓		✓
Electoral controls		✓		✓		✓		✓		✓

Notes: Each entry is estimated separately. Panel A and B present estimates of equation 3.1 and panel D of equation 3.2. First-stage estimates uses an inverse logarithmic decay function of the distance to the closest entry port. Odd Columns do not include any control and even columns include all the control variables. Dependent variables in Columns (1) to (8) are from the 2015 mayoral election and in Columns (9) and (10) are from the 2016 peace agreement plebiscite, as reported in the column header. Robust standard errors presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table 3.8: Relative importance of control variables and the peace plebiscite results

Mediator	Coefficient (1)	Standard Error (2)
Peace Plebiscite	0.024***	0.004
Homicides	-0.000	0.000
Robbery	0.000	0.000
Right-wing mayor	-0.000	0.000
Left-wing mayor	-0.001	0.001
Total Explained	0.023***	0.004

Notes: Gelbach (2016) decomposition of the role of the various mechanisms for explaining the votes for the *Centro Democrático* candidate, Iván Duque. "Total Explained" represents the difference in the estimate of the variable "Venezuelan migrants" between the full and the base models. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Appendix A: Appendices for Chapter 1

A.1 Conceptual framework model

I guide my conceptual framework using the model on Hirshleifer (1991) and extended by Hirshleifer (2001) and Dixit (2004). There are similar models for producer and grabbers that also depend on other characteristics, such as the quality of institutions (Mehlum, Moene and Torvik, 2006).

The simplest economy of the model is a one-period static model with two participants, $i = 1, 2$, say a farmer and a cartel.[1] Each participant is endowed with one unit of resource, which may be her own labor time or some other productive asset. Each participant can allocate the resource into production effort Y_i or conflict effort by either defending her output against the other's aggression Z_i or in attacking the other's output X_i. Note, conflict here is the mere failure to cooperate, therefore the decision to allocate the resource to production entails the possibility of trade. The resource constraint is:

$$Y_i + Z_i + X_i \leq 1$$

.

[1] One could think on two groups of participants whose actions are coordinated, such that the two participants could also be a representative member of their respective group.

Each participant seek to maximize its expected payoff generated by the two sides' productive efforts. The resource allocated to production yield an output $(Y_i)^\alpha$ where the parameter α is positive and indicates the return to scale in production. The probability of keeping each participant own share of production depend upon the conflict efforts given by

$$p_i = \frac{Z_i^\beta}{Z_i^\beta + \theta X_j^\beta}$$

,

where β is a parameter between 0 and 1 that represents the returns to scale to conflict, and the parameter θ measures the efficacy of the offense relative to the other participant j defense. The expected payoff, or aggregated income, is:

$$\Pi_i = Y_i^\alpha \frac{Z_i^\beta}{Z_i^\beta + \theta X_j^\beta} + Y_j^\alpha \frac{\theta X_i^\beta}{Z_j^\beta + \theta X_i^\beta}$$

,

for $i = 1, 2$ and $j \neq i$. This specification involves two assumptions: (i) income is jointly produced and falls into a common pool or resources available for capture, and (ii) conflict activity does not damage the resource or the production process. Note, if neither participant decide to attack, each participant's payoff will solely depend on her own production.

To solve the equilibrium of this game, one could assume a symmetric outcome where the two participants choose identical allocations (Dixit, 2004). However, this assumption would not be real in the present case of a farmer and a cartel as farmers might not be interested in getting the production of the cartel or attacking them. Nonetheless, farmers could also adopt self-defense groups or vigilantism as means of protecting themselves from organized crime (Asfura-Heim and

Espach, 2013; Rios, 2013). Furthermore, the setting implies a simultaneous decision, whereas it is possible that the farmer react to the cartel's decision (or vice versa). This type of asymmetric solutions are studied by Hirshleifer (2001), including the possibility of the Stackelberg situation where there is a "leader" that chooses first the conflict intensity.

To simplify the analysis, I assume a symmetric equilibrium. Under this assumption, the Nash equilibrium of this game can be found after some algebra. The equilibrium quantities are:

$$Y_i = \frac{\alpha(1+\theta)}{\alpha(1+\theta)+2\beta\theta}$$

$$Z_i = X_i = \frac{\beta\theta}{\alpha(1+\theta)+2\beta\theta}$$

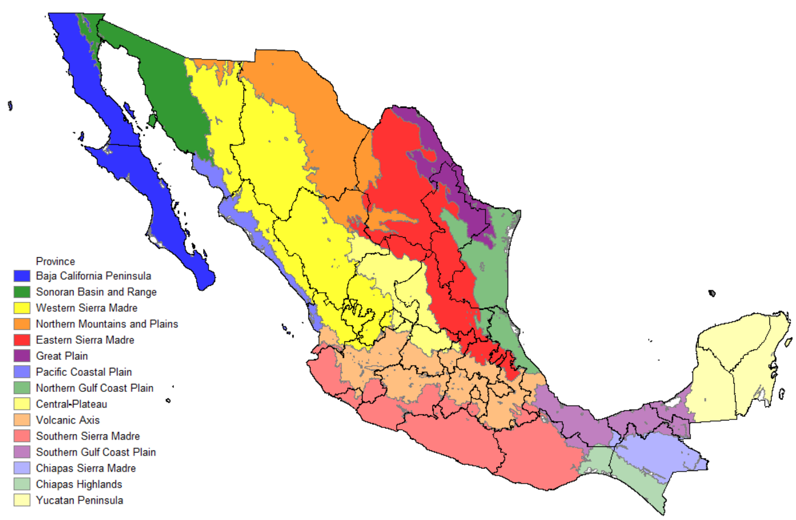
This equilibrium is not economically efficient, as the ideal equilibrium would be for neither player to spend anything on aggression (X_i). However, this possibility would not be a Nash equilibrium as deviations would give the participant a higher payoff. Similar to the prisoner's dilemma, if I decide to spend zero on defense, then the other participant could take all my production by spending just some units on her resource on attacking.

Intuitively, the equilibrium allocation of the productive use of resource (Y_i) changes for each participant positively if there is an improvement in the technology of production (α), as the use of one unit in production would yield more output. On the other hand, if the returns to fighting (β) improves, the productive use of resources would decrease as now fighting give more returns. Similarly, if the technology of fighting favors defense over offense (θ), then fighting will yield a higher payoff.

A.2 Physiographic regions in Mexico

Mexico is divided into 15 physiographic regions based on geological characteristics, presented in Figure A.1. These regions do not overlap with political or administrative divisions, as shown in the black line divisions. Note, there are states that belong to only one physiographic region, while there are others that have more than one. For instance, the states of Quintana Roo and Yucatán, located in the eastern-most part of the country, belong only to the region of Yucatan Peninsula; while the states Nuevo León and Tamaulipas in the north-eastern part of the country, belong to three physiographic regions: Great Plains, Eastern Sierra Madre, and Northern Gulf Coastal Plain.

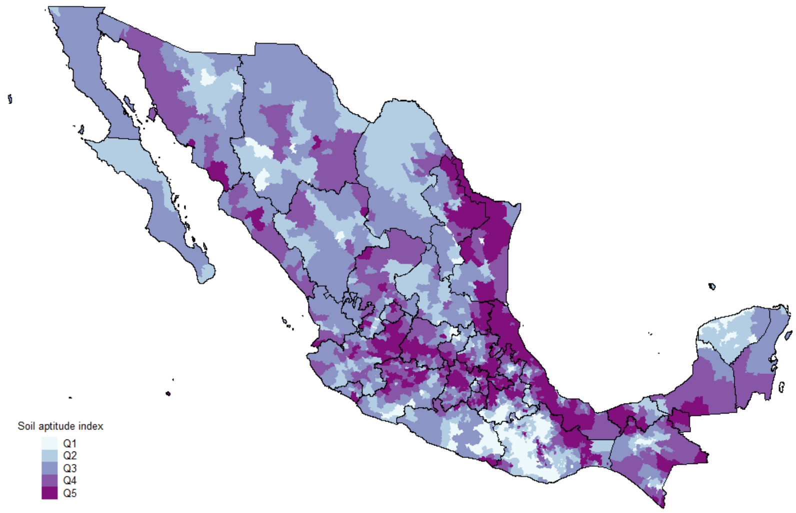
Figure A.1: Mexico's physiographic regions and state division



Source: data from INEGI. The figure represents the 15 physiographic regions of Mexico. Black lines represent the states.

Soil quality for growing agricultural crops is reported in Figure A.2. This index summarizes the areas which are more apt to be used for agricultural purposes.

Figure A.2: Soil index for agricultural crops



Source: data from INEGI. Soil index for growing agricultural crops. Darker shades represent soils that are more apt to use for agricultural purposes.

A.3 Classification of homicides

A.3.1 Classification of death certificates as homicides

I use data of the death certificates from the National Department of Health Information (SINAIS) to classify deaths as homicides. This classification follows the list of causes and corresponding ICD-10 codes from the World Health Organization (WHO). Homicides are classified under codes X85 to Y09, that correspond to Assault (homicide), and its sequalae, under code Y871. These correspond to 209 different causes; however, the majority correspond to "assault by other and unspecified firearm discharge, street and highway." The top 10 causes represents almost 80% of the total cases and the top 2 causes represents around 92% of them. Table A.1 presents the code and description of the top 20 cases, the count and the percentage with respect to the total number of cases.

Table A.1: Summary and description of top 20 causes of homicides

Rank	Code	Description	Count	Percent (%)	Cum. (%)
1	X954	Assault by other and unspecified firearm discharge, street and highway	131,430	39.14	39.14
2	X959	Assault by other and unspecified firearm discharge, unspecified place	25,094	7.47	46.61
3	X950	Assault by other and unspecified firearm discharge, home	22,285	6.64	53.24
4	X994	Assault by sharp object, street and highway	20,425	6.08	59.33
5	X958	Assault by other and unspecified firearm discharge, other specified places	18,795	5.6	64.92
6	Y094	Assault by unspecified means, street and highway	16,818	5.01	69.93
7	X914	Assault by hanging, strangulation and suffocation, street and highway	8,547	2.55	72.48
8	X990	Assault by sharp object, home	8,531	2.54	75.02
9	Y099	Assault by unspecified means, unspecified place	7,817	2.33	77.34
10	X957	Assault by other and unspecified firearm discharge, farm	6,814	2.03	79.37
11	X999	Assault by sharp object, unspecified place	6,400	1.91	81.28
12	X955	Assault by other and unspecified firearm discharge, trade and service area	6,289	1.87	83.15
13	X910	Assault by hanging, strangulation and suffocation, home	5,517	1.64	84.79
14	Y098	Assault by unspecified means, other specified places	4,926	1.47	86.26
15	Y090	Assault by unspecified means, home	4,779	1.42	87.68
16	X998	Assault by sharp object, other specified places	4,634	1.38	89.06
17	X919	Assault by hanging, strangulation and suffocation, unspecified place	3,969	1.18	90.25
18	X918	Assault by hanging, strangulation and suffocation, other specified places	2,773	0.83	91.07
19	Y004	Assault by blunt object, street and highway	2,257	0.67	91.74
20	X934	Assault by handgun discharge, street and highway	1,751	0.52	92.27
		All other causes of homicides	25,975	7.73	100
		Total number of homicides	335,826	100.0	

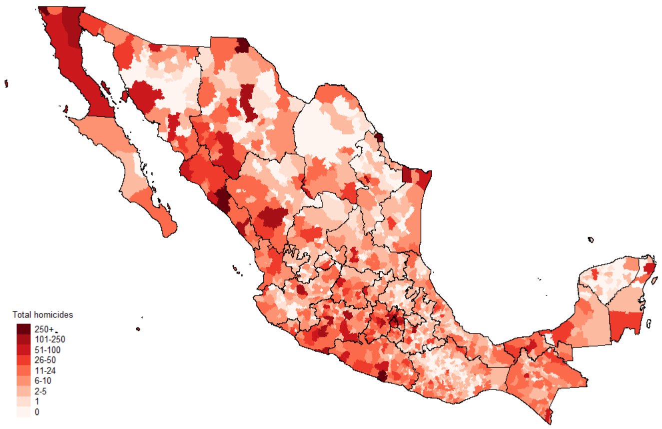
Using this information, I calculate the homicide rate for each year both as a unique value and as an average of the homicide rate for each municipality. Table A.2 presents these estimates. Note, there is a lot of variation in homicide rate. For instance, cities like Ciudad Juarez exhibited homicide rates of 216 victims per 100,000 inhabitants in 2010, a casualty rate that is comparable to that of war zones (Rios, 2013, p. 139). Figure A.3 presents the spatial distribution of the total number of homicides. Former president Felipe Calderon started the Mexican war on drugs at the end of 2006. President Enrique Peña Nieto, his successor, attempted to curve down the increase in conflict by sending national forces to take control back of the territories rather than relying on local police (Althaus and Dudley, 2014).

Table A.2: Homicide rate by year

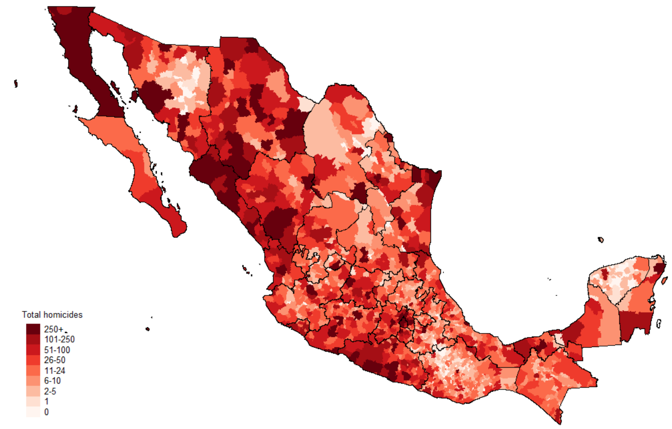
	National		Municipal data	
Year	Homicides	rate	Mean rate	S.D.
2,000	10,737	8.52	11.05	24.06
2,001	10,285	8.16	10.30	23.08
2,002	10,088	8.01	10.28	25.25
2,003	10,087	8.00	10.17	23.19
2,004	9,329	7.40	10.03	26.41
2,005	9,921	7.87	9.63	24.56
2,006	10,452	8.29	10.56	26.32
2,007	8,867	7.04	9.77	26.43
2,008	14,006	11.11	12.03	30.40
2,009	19,803	15.71	15.89	51.11
2,010	25,757	20.44	22.30	78.82
2,011	27,213	21.60	22.41	68.00
2,012	25,967	20.61	23.25	62.07
2,013	23,063	18.30	20.48	42.89
2,014	20,010	15.88	17.35	34.84
2,015	20,762	16.48	17.31	37.36
2,016	24,559	19.49	17.11	29.09
2,017	32,079	25.46	21.03	36.43
2,018	36,685	29.11	22.59	46.43
2,019	36,661	29.09	22.84	40.63
2,020	36,773	29.18	22.27	51.93

Note: Own calculation using homicide data coming from SINAIS. Rates are per 100,000 people and are calculated using population in 2020 census. National homicide rate is calculated dividing the number of homicides in the population. Homicide rate at the municipal level is calculated using the number of cases in each municipality and averaging them by year.

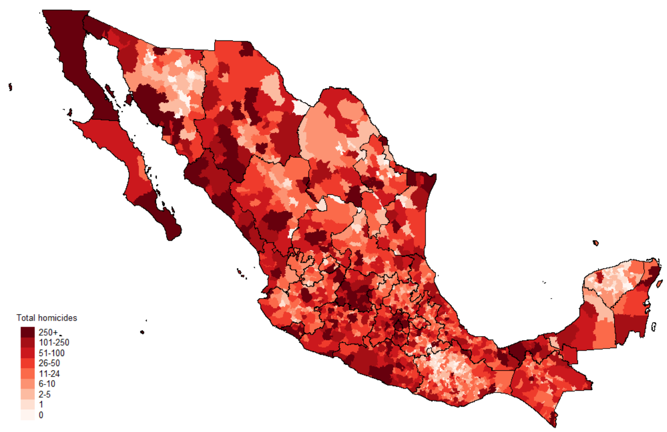
Figure A.3: Evolution and distribution of homicides in Mexico



(a) 2005 to 2006



(b) 2007 to 2012

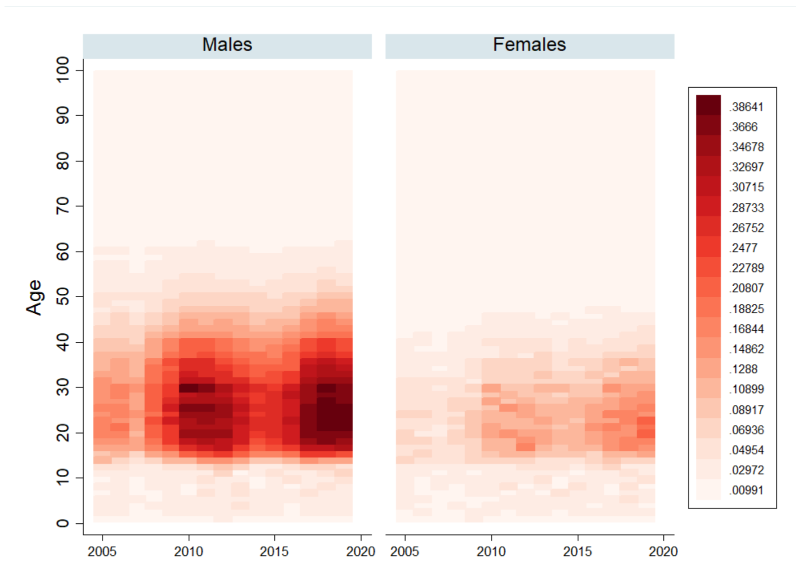


(c) 2013 to 2019

Note: Own calculation using homicide data coming from SINAIS. Total number of homicides from 2005 to 2006 presented in Panel a; from 2007 to 2012 in Panel b; and from 20013 to 2019 in Panel c. Former president Felipe Calderon started the Mexican war on drugs started at the end of 2006 and his presidential term ended at the end of 2012.

My classification of homicides based on victims' characteristics rely on the assumption that victims that belong to the group of males aged between 15 and 40 are more likely to join criminal activities. I provide some evidence in Figure A.4 that shows that the majority of homicides are concentrated in males between the ages 15 and 40 consistently for every year in my sample. Similarly, Figure A.5 shows that males aged between 15 and 40 was the group of the population more affected by the Mexican war on drugs.

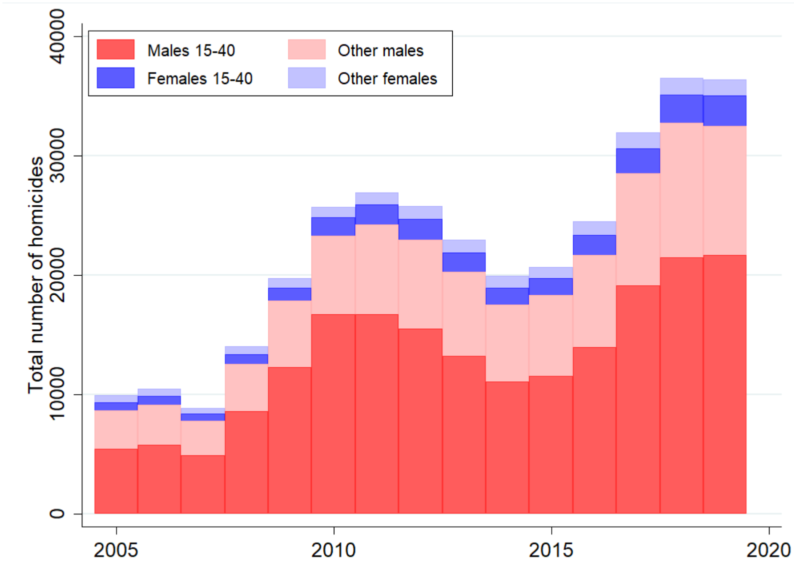
Figure A.4: Distribution of homicides by age and gender over time



Note: Own calculation using homicide data coming from SINAIS.

I have two sources of information about homicides: death certificates and police reports. Although the two figures are not the same in terms of total number of cases, they follow each other closely. For instance, the total number of deaths between 2011 and 2019, the period in common, is 2,207 coming from death certificates and 2,392 coming from police reports. This is not an uncommon phenomenon between sources of information, as described by Atuesta and Pérez-Dávila (2017). Figure A.6 presents in the left panel the observations at the municipality-

Figure A.5: Distribution of homicides by age groups and gender over time

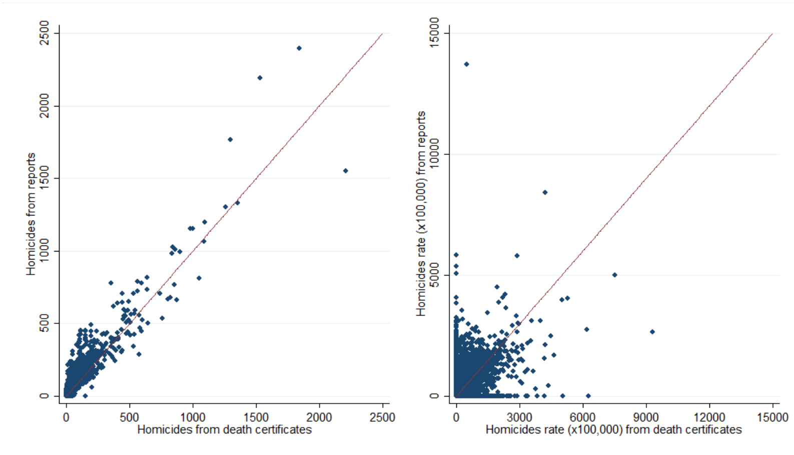


Note: Own calculation using homicide data coming from SINAIS. "Males 15-40" and "Females 15-40" count the number of male victims between the ages of 15 and 40, and the number of females between the same age group, respectively. "Other males" and "Other females" count the number of male victims outside the age group of 15 to 40, and the number of female victims outside that age group, respectively.

year level. The majority follow the 45-degree line, suggesting that they are the same. However, there is more noise when comparing their respective homicide rates (Figure A.6, right panel).

A possible explanation for this discrepancy is that they are not registered in the same place of occurrence. Indeed, death certificates include three separate locations: occurrence, registration, and residence. The place of registration is where the death certificate is generated, the place of occurrence is where the death happened, and the place of residence is where the individual deceased had her usual residence. These three locations may differ one from the other if the location where the death happened is different from the usual location of the residence. There is almost perfect correlation between the place of residence and occurrence, yet these two are 70 percent correlated with the municipality where the death was registered.

Figure A.6: Comparison of homicide rates between sources of information



Note: Panel on the left compares the homicides from the two sources of information. Panel on the right presents the comparison in homicide rates.

A.4 Additional robustness

A.4.1 Alternative explanatory variables

In this section I explore differences in my measure of potential production. In particular, my measure is defined as:

$$Z_{mt} = \hat{S}_m \times \log(D_t)$$

First, I present alternative measures for the share part $\hat{S}_m$, then I present alternative measures for the shift part $\log(D_t)$, with and without the logarithm.

A.4.1.1 Alternative sources of cross-sectional variation

One alternative to my measure of areas suited to absorb the demand for avocado is to use different measures of the avocado suitability index (see Section 1.5.1.1) or to use the area devoted to avocado production to the initial period (Ferri, 2022; Goldsmith-Pinkham, Sorkin and Swift, 2020; Jaeger, Ruist and Stuhler, 2018). Specifically, I use the share of agricultural area devoted to avocado production in 2003, the first year I observe information at this level. To ease comparisons, I interact the new measures with the (log of the) US demand and standardize them to have zero mean and unit variance. My estimates are consistent across these different measures, in terms of sign and significance, including the share of area devoted to avocado production in 2003 (Table A.3). Differences in point estimates come from the variability of the suitability indices as they are no longer bound between zero and one, implying that a one-standard-deviation increase does not represent the same changes in the potential production.

Table A.3: Robustness using different measures of the avocado suitability index

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable: Homicide rate agricultural workers									
Suitability × log(US demand)	1.612***	1.060**								
	(0.428)	(0.452)								
Suitability (OLS)× log(US demand)			2.151***	1.538***						
			(0.393)	(0.441)						
Suitability (yield)× log(US demand)					2.293***	1.779***				
					(0.409)	(0.418)				
Suitability (IHS yield)× log(US demand)							2.322***	1.744***		
							(0.393)	(0.423)		
Share avocados 2003× log(US demand)									0.538**	0.449
									(0.246)	(0.274)
Observations	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529
Number of municipalities	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438
Mean Dep. Var.	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized in all specifications to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The dependent variable is the homicide rate of agricultural workers per 100,000 people and the bottom panel reports its mean in the sample. Suitability measures 1 to 4 come from the predictions of Table 1.3, in the same order. Share of avocados in 2003 is the share of cultivated land devoted to avocados in 2003. Block bootstrap standard error using clusters at the municipal level and 200 replications are reported parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

A.4.1.2 Alternative sources of temporal variation

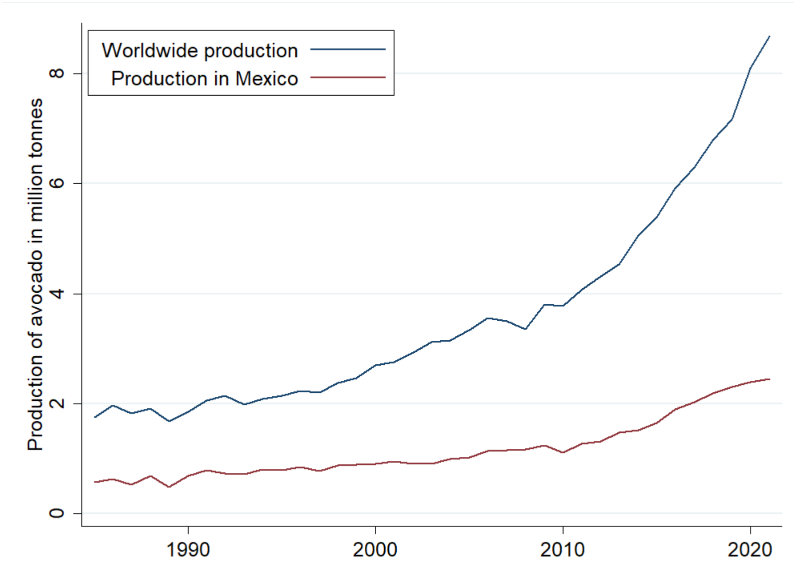
The shock factor in my measure of the areas suited to absorb the demand for avocado relies on the United States demand for the crop, which I argue is exogenous to the conflict intensity (see Section 1.4.2). Alternatively, I proxy the exogenous demand shock for Mexican avocados using the growth of its industry in Mexico (Figure A.7), to avoid the endogeneity coming from the idiosyncratic industry-location component (Adão, Kolesár and Morales, 2019; Borusyak, Hull and Jaravel, 2021; Goldsmith-Pinkham, Sorkin and Swift, 2020); the global production of avocados excluding Mexico;[2] or using the interest in avocados by looking at the number of searches on Google Trends to proxy for demand (Estancona and Tiscornia, 2022). Specifically, I use the number of searches for "Avocado Toast" in the United States and for "Avocado" in the world (Figure A.8).

I re-estimate equation 1.1 using these measures interacted with the avocado suitability index and present my results in Table A.4. My estimates are consistent across specifications in terms of sign and significance. Similar to the alternative suitability indices, the difference in point significance comes from the difference in volume of avocado production. Similarly, I create measures of potential production interacting the avocado suitability index and the measures of demand shock in levels, i.e., without taking the log. My results are robust to these alternative explanatory variables (Table A.5) although, the point estimates are smaller compared to their version in logs. The only exception is the measure Google searches for "avocado toast" in the

[2] I assume that the interest in avocados, measured as growth in production, follows a similar trend to the interest in Mexican avocados. Using the global production excluding Mexico mimics the work of Dube and Vargas (2013) who examine changes in prices of coffee in Colombia by inspecting changes in the three other leading coffee exporting nations. Figure A.7 presents the world and the Mexican avocado production, and resonates with Huang et al. (2023), who discusses the increasing global demand for avocado consumption.

United States, which is not statistically different from zero. A possible explanation for this result is that I am restricting the search to one use of avocados, yet it might not be the more demanded.[3]

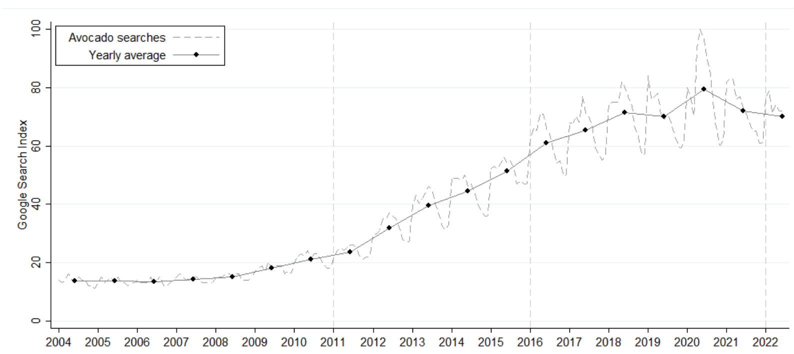
Figure A.7: Production of avocado in the world and in Mexico



Note: Source: the Food and Agriculture Organization (FAO). Blue line represents the total avocado production in the world. Red line represents the production of avocado in Mexico. The difference between the two is the production of avocado in the world without Mexico.

[3]For instance, guacamole is consumed extensively during Super Bowl Sunday and *Cinco de Mayo*, the Mexican holiday that is more widely celebrated in the United States. Avocado consumption these days is 100 million pounds and 90 million pounds of avocado respectively (Miller, 2020). Other uses include avocado oil, smoothies, or as part of salads.

Figure A.8: Global importance of avocado using Google search trends



Note: Source: Google Trends. The Google Search Index (vertical axis) is a normalized index of interest respect to the highest point (value 100). It is a measure of relative importance. Dashed vertical lines represents improvements applied on location and data collection. Dots represent the average value of the index calculated in a calendar year.

Table A.4: Robustness using different measures of growth in the demand for avocado

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable: Homicide rate agricultural workers									
Suitability × log(US demand)	1.612***	1.060**								
	(0.428)	(0.452)								
Suitability × log(Mexican exports)			8.859***	5.623**						
			(2.661)	(2.771)						
Suitability × log(World production)					14.309***	8.524*				
					(4.273)	(4.458)				
Suitability × log(Google avocado)							1.565***	1.035**		
							(0.459)	(0.483)		
Suitability × log(Google toast in US)									0.574***	0.297*
									(0.160)	(0.168)
Observations	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	34,344	34,095
Number of municipalities	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438
Mean Dep. Var.	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized in all specifications to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The dependent variable is the homicide rate of agricultural workers per 100,000 people and the bottom panel reports its mean in the sample. The explanatory variable is the interaction between avocado suitability index and the log of the measure of demand for avocado. US demand is the demand for avocado in the United States; Mexican exports is the total production of avocado for export; world production is the total production of avocados in the world excluding that of Mexico; google avocado is the google search index for the word avocado in the world; and google toast in US is the google search index for the word avocado toast in the US. Block bootstrap standard error using clusters at the municipal level and 200 replications are reported parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table A.5: Robustness using different measures of demand for avocado

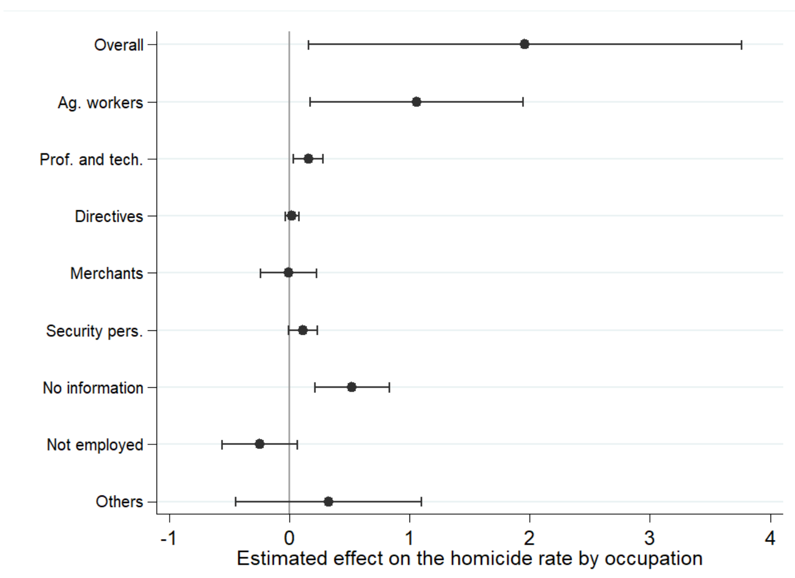
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable: Homicide rate agricultural workers									
Suitability × US demand	0.855***	0.576**								
	(0.282)	(0.292)								
Suitability × Mexican exports			4.608***	2.974*						
			(1.736)	(1.749)						
Suitability × World production					7.434***	4.520				
					(2.754)	(2.812)				
Suitability × Google avocado							0.825***	0.553*		
							(0.303)	(0.310)		
Suitability × Google toast in US									0.313***	0.161
									(0.117)	(0.122)
Observations	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	34,344	34,095
Number of municipalities	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438	2,456	2,438
Mean Dep. Var.	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779	5.779
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized in all specifications to have zero mean and unit variance. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The dependent variable is the homicide rate of agricultural workers per 100,000 people and the bottom panel reports its mean in the sample. The explanatory variable is the interaction between avocado suitability index and the measure of demand for avocado. US demand is the demand for avocado in the United States; Mexican exports is the total production of avocado for export; world production is the total production of avocados in the world excluding that of Mexico; google avocado is the google search index for the word avocado in the world; and google toast in US is the google search index for the word avocado toast in the US. Block bootstrap standard error using clusters at the municipal level and 200 replications are reported parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

A.4.2 Homicides by occupation reported

Using information from death certificates, I classify the homicides into three groups: agricultural workers, (possibly) cartel related, and rest of the population. In this section, I use the information on the victims' occupation to show that the change in the overall homicide rate is explained by changes in the homicide rate of agricultural workers and not of other types of workers. Specifically, I show for professionals and technicians, directives, merchants, security personnel, and for those that there is no information specified and for those that are not working. Given a change in the classification of occupations in 2013, these were the only groups that are comparable. I put together the rest of the occupations and show them in a category of their own. Figure A.9 presents the estimates for these groups and Table A.6 reports them.

Figure A.9: Effects on homicide rate by victims' occupation



Note: Markers correspond to estimates of equation 1.1 where the explanatory variable are the homicide rates for the occupations reported in the vertical axis. All estimates include municipality and time fixed effects, and the interaction of the municipal-level controls and year fixed effects. Block bootstrap standard error using clusters at the municipal level and 200 replications. Horizontal lines represent 95% confidence intervals.

Table A.6: Effects on homicide rate by victims' occupation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Homicide rate by occupation							
	Homicides	Agricultural workers	Professional & technicians	Directives	Merchants	Security personnel	Without information	Not employed	other
Suitability × log(US demand)	1.960**	1.060**	0.159***	0.025	-0.003	0.112*	0.524***	-0.247	0.329
	(0.918)	(0.452)	(0.062)	(0.029)	(0.118)	(0.061)	(0.157)	(0.160)	(0.394)
Observations	36,529	36,529	36,529	36,529	36,529	36,529	36,529	36,529	36,529
Number of municipalities	2,438	2,438	2,438	2,438	2,438	2,438	2,438	2,438	2,438
Mean Dep. Var.	16.927	5.758	0.425	0.083	1.187	0.439	1.118	2.056	5.862
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized. All columns include municipality and year fixed effects, and the interaction of the municipal-level controls and year fixed effects. The column header reports the occupation and the bottom panel reports its mean in the sample. All the outcomes are rates per 100,000 people. Block bootstrap standard error using clusters at the municipal level and 200 replications are reported parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

A.4.3 Alternative populations of study

To rule out the role of very small municipalities or big cities, I exclude from my sample municipalities above the 95th and below the 5th percentile. Table A.7 reports these estimates and show that they are robust to this exclusion. In addition, I show in Table A.8 that my results are robust to excluding the municipalities whose number of homicides is above the 99th percentile, to account for potential outliers in the reports.

Table A.7: Effects of potential to absorb demand for avocado on homicides excluding small and big municipalities

	(1) Homicides	(2) Homicides	(3) Agricultural worker	(4) Agricultural worker	(5) Cartel related	(6) Cartel related	(7) Rest of population	(8) Rest of population
Suitability × log(US demand)	0.707	1.361	1.418***	0.889*	-0.241	0.314	-0.470	0.158
	(0.911)	(1.011)	(0.429)	(0.475)	(0.434)	(0.432)	(0.338)	(0.354)
Observations	33,141	32,876	33,141	32,876	33,141	32,876	33,141	32,876
Number of municipalities	2,210	2,192	2,210	2,192	2,210	2,192	2,210	2,192
Mean Dep. Var.	16.951	16.919	5.950	5.928	5.887	5.883	5.113	5.107
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. All the outcomes are rates per 100,000 people. Cartel related refers to non-agricultural male workers between the ages of 15 and 40. Municipalities below the 5th percentile and above the 95th percentile of the population are excluded. Block bootstrap standard error using clusters at the municipal level and 200 replications are reported parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table A.8: Effects of potential to absorb demand for avocado on homicides excluding potential outliers in levels

	(1) Homicides	(2) Homicides	(3) Agricultural worker	(4) Agricultural worker	(5) Cartel related	(6) Cartel related	(7) Rest of population	(8) Rest of population
Suitability × log(US demand)	0.885	1.593*	1.462***	0.977**	-0.271	0.352	-0.379	0.235
	(0.894)	(0.916)	(0.342)	(0.409)	(0.434)	(0.426)	(0.359)	(0.356)
Observations	36,425	36,160	36,392	36,127	36,425	36,160	36,423	36,158
Number of municipalities	2,452	2,434	2,455	2,437	2,452	2,434	2,452	2,434
Mean Dep. Var.	16.576	16.545	5.423	5.398	5.736	5.731	5.050	5.044
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable and the bottom panel reports its mean in the sample. All the outcomes are rates per 100,000 people. Observations in the top 1st percentile of the outcome in levels are excluded. Cartel related refers to non-agricultural male workers between the ages of 15 and 40. Block bootstrap standard error using clusters at the municipal level and 200 replications are reported parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

A.5 Additional results

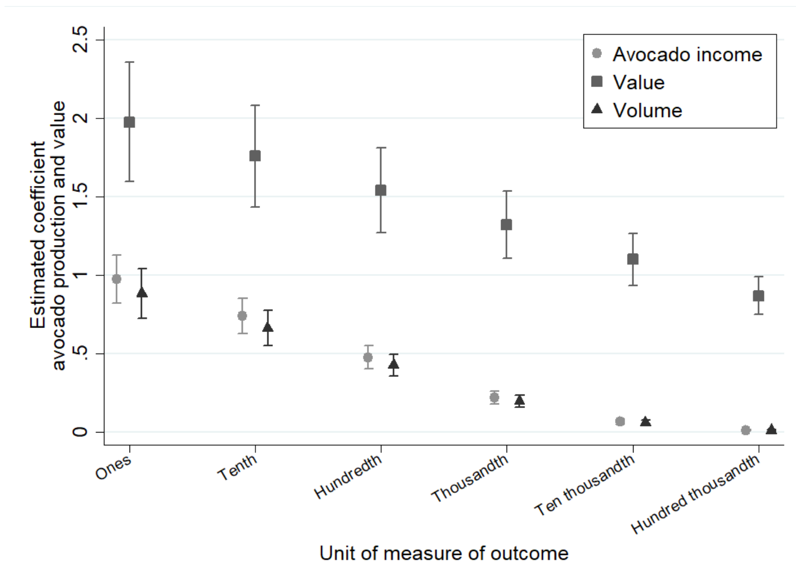
A.5.1 Sensitivity to functional form of avocado production

Table 1.5 presents in the last 4 columns the estimates on avocado production taking the inverse hyperbolic sine (I.H.S.), as opposed to taking the natural log given the presence of zeros. However, the I.H.S. transformation yields different estimations depending on the units of measure of the outcome. In particular, the transformation applied to large units for the outcome variable, say production in millions, will yield a distribution that will resemble a discrete variable, as the I.H.S. transformation separates the zeros from the positive values. On the other hand, small units for the outcome will yield a continuous variable where the is no separation between the zeros and other values and whose interpretation is closer to a continuous variable (see McKenzie (2023) for a detailed discussion). Figure A.10 replicate the estimates of the last four columns of Table 1.5 and show that they are not sensitive to the unit of measure. Notice, values to the left of the horizontal axis resemble a discrete variable and captures the extensive margin; while values to the right resemble a continuous variable and captures the intensive margin.

A.5.2 Additional results on the role of competition and cooperation

Tables A.9 and A.10 present, respectively, the heterogeneous effects by presence of cartels in the previous period and by presence of multiple cartels in the previous period using as outcomes the overall homicide rate and the homicide rate of the rest of the population.

Figure A.10: Robustness to functional form of actual production



Note: Markers correspond to estimates of equation 1.1 using as dependent variable the inverse hyperbolic sine (IHS) of the avocado income per capita (circles), the value of avocado production (squares), and the volume of avocado production (triangles), divided by the unit reported in the horizontal axis. All estimates include municipality and time fixed effects, and further control for the interaction of the municipal-level controls and year fixed effects. Vertical lines represent 95% confidence intervals.

Table A.9: Heterogeneous effects on homicides by cartel presence in the previous period

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Dep. var: Overall homicide rate						Dep. var: rest of the population					
	any	any	traditional	traditional	non-traditional	non-traditional	any	any	traditional	traditional	non-traditional	non-traditional
Suitability × US demand	1.333	2.694**	1.568	2.853**	-0.085	1.003	-0.308	0.626	-0.249	0.652	-0.753**	0.021
	(1.087)	(1.170)	(1.074)	(1.153)	(0.936)	(1.004)	(0.431)	(0.472)	(0.426)	(0.458)	(0.370)	(0.411)
Suitability × US demand × cartel_$t-1$	0.056	-0.766	-0.219	-0.996**	2.799***	1.809***	0.042	-0.333*	-0.025	-0.384**	0.920***	0.544***
	(0.498)	(0.476)	(0.503)	(0.470)	(0.662)	(0.570)	(0.196)	(0.192)	(0.195)	(0.186)	(0.209)	(0.197)
Observations	36,734	36,469	36,734	36,469	36,734	36,469	36,734	36,469	36,734	36,469	36,734	36,469
Number of municipalities	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434	2,452	2,434
Mean Dep. Var.	16.956	16.956	16.956	16.956	16.956	16.956	5.177	5.177	5.177	5.177	5.177	5.177
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable, the type of cartel, and the bottom panel reports its mean in the sample. In Columns (1) to (6) the dependent variable is the overall homicide rate in the municipality, and in Columns (7) to (12) is the homicide rate of the population that is not agricultural worker nor potentially cartel related, i.e., non-agricultural male workers between the ages of 15 and 40. All the outcomes are rates per 100,000 people. $cartel_{t-1}$ is a dummy variable that takes value of 1 if there was a cartel present in the previous years. Block bootstrap standard error using clusters at the municipal level and 200 replications are reported parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table A.10: Heterogeneous effects on homicides by cartel presence of multiple cartels in the previous period

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Dep. var: Overall homicide rate						Dep. var: rest of the population					
	any	any	traditional	traditional	non-traditional	non-traditional	any	any	traditional	traditional	non-traditional	non-traditional
Suitability × US demand	0.541	1.945*	0.981	2.282**	-0.207	0.894	-0.549	0.509	-0.376	0.626	-0.917**	-0.091
	(1.027)	(1.000)	(1.022)	(0.980)	(0.922)	(0.917)	(0.507)	(0.476)	(0.503)	(0.467)	(0.458)	(0.437)
Suitability × US demand × one cartel_$t-1$	0.564	-0.083	0.092	-0.493	1.751***	0.988	0.418*	0.003	0.284	-0.109	1.013***	0.596**
	(0.560)	(0.523)	(0.537)	(0.522)	(0.628)	(0.604)	(0.233)	(0.219)	(0.225)	(0.207)	(0.295)	(0.295)
Suitability × US demand × two+ cartel_$t-1$	1.049*	0.055	0.630	-0.296	5.107***	3.623***	0.761***	0.138	0.611**	0.017	3.216***	2.412***
	(0.571)	(0.586)	(0.583)	(0.571)	(1.065)	(1.129)	(0.278)	(0.276)	(0.295)	(0.281)	(0.582)	(0.596)
Observations	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529	36,794	36,529
Number of municipalities												
Mean Dep. Var.	16.956	16.956	16.956	16.956	16.956	16.956	5.999	5.999	5.999	5.999	5.999	5.999
Municipality FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls × Year Fe		✓		✓		✓		✓		✓		✓

Note: Each column reports estimates of equation 1.1. Potential to absorb demand for avocado variable is standardized. All columns include municipality and year fixed effects. Even columns further control for the interaction of the municipal-level controls and year fixed effects. The column header reports the dependent variable, the type of cartel, and the bottom panel reports its mean in the sample. In Columns (1) to (6) the dependent variable is the overall homicide rate in the municipality, and in Columns (7) to (12) is the homicide rate of the population that is not agricultural worker nor potentially cartel related, i.e., non-agricultural male workers between the ages of 15 and 40. All the outcomes are rates per 100,000 people. $onecartel_{t-1}$ is a dummy variable that takes value of 1 if there was only one cartel present in the previous year, and $two + cartel_{t-1}$ is a dummy that takes the value of 1 if there were two or more cartels present in the previous year. Block bootstrap standard error using clusters at the municipal level and 200 replications are reported parentheses. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Appendix B: Appendices for Chapter 3

B.1 Data and additional results

Summary statistics – Table B.1 presents the summary statistics of the main variables used in the analysis. Panel A presents the main outcomes, Panel B the migration measure and the main instrument measure before standardizing. Panel C through Panel E presents the statistics for the control variables grouping them as presented in the main tables. Table B.2 presents the summary statistics for the additional variables used in the analysis. Panel A summarizes the different decay functions; Panel B, the migration measures; Panel C, the election results from the second round of the 2018 presidential election; Panel D, the first round election results for each candidate in the 2018 presidential election; Panel E, the results of the 2014 presidential election aggregated by ideology; and Panel F, the two outcomes based that summarizes the political preferences for the Centro Democratico party.

First stage results – Figure B.1 presents the IV estimates using different exponential decay functions. Figure B.1a presents the estimates of equation 3.1 or the second stage, while Figure B.1b presents estimates of equation 3.2. The exponential decay function takes the form

$$f(distance) = \exp\left\{-\frac{1}{\lambda}distance\right\}$$

Table B.1: Descriptive statistics for the main variables

	Obs. (1)	Mean (2)	Std. Dev. (3)	Min (4)	Max (5)
Panel A: Election outcomes in 2018 first-round					
Votes for Duque	1,122	0.469	0.185	0.000	0.871
Votes for right	1,122	0.576	0.175	0.064	0.924
Votes for center	1,122	0.150	0.097	0.012	0.535
Votes for left	1,122	0.237	0.184	0.011	0.879
Turnout	1,122	0.499	0.100	0.066	0.753
Panel B: Venezuelan presence and instrument					
Shock Venezuelan migrants	1,122	0.474	0.862	0.000	7.335
Inverse logaritmic distance to 5 entry ports	1,122	0.155	0.022	0.129	0.509
Panel C: Border controls					
Border with brasil	1,122	0.010	0.099	0.000	1.000
Border with ecuador	1,122	0.009	0.094	0.000	1.000
Border with panama	1,122	0.004	0.060	0.000	1.000
Border with peru	1,122	0.006	0.079	0.000	1.000
Border with venezuela	1,122	0.037	0.188	0.000	1.000
Panel D: Distance controls					
Distance to Bogota	1,122	321.6	194.6	0.0	1,270.9
Distance to Department's capital	1,122	81.5	60.6	0.0	493.1
Distance to closest principal market	1,122	130.0	111.6	0.0	926.5
Panel E: Administrative controls					
Area (km2)	1,122	1,018	3,201	15	65,674
Elevation (MASL)	1,122	1,140	1,156	2	25,221
Population census (2005)	1,113	36,983	230,425	4	6,740,859
Rurality index (2005)	1,118	0.589	0.241	0.002	1.000
Migrant population (2018)	1,122	0.129	0.249	0.000	4.426
Panel F: Mayoral election and peace plebiscite controls					
Votes for NO option	1,122	0.460	0.167	0.000	0.821
Votes for center	1,035	0.597	0.255	0.000	0.992
Votes for left	1,035	0.085	0.158	0.000	0.985
Votes for right	1,035	0.241	0.223	0.000	0.987
Left-wing mayor	1,035	0.060	0.237	0.000	1.000
Right-wing mayor	1,035	0.213	0.409	0.000	1.000
Center-wing mayor	1,035	0.726	0.446	0.000	1.000

Notes: The table present the summary statistics for the main control variables by its classification. The observations are at the municipal level.

Table B.2: Descriptive statistics for the additional variables

	Obs. (1)	Mean (2)	Std. Dev. (3)	Min (4)	Max (5)
Panel A: Decay functions					
Inverse logaritmic distance to 5 entry ports	1,122	0.155	0.022	0.129	0.509
Inverse linear distance to 5 entry ports	1,122	0.004	0.010	0.001	0.286
Min. distance to 5 entry points	1,122	414.7	229.6	3.5	1,135.2
Inverse logaritmic distance to Cucuta	1,122	0.153	0.018	0.127	0.318
Distance to Cucuta entry port	1,122	432.3	228.0	11.6	1,285.1
Panel B: Migration measures					
Presence Venezuelan migrants	1,122	0.474	0.862	0.000	7.335
Presence Venezuelan migrants $<$1 year	1,122	0.370	0.546	0.000	4.171
Presence Venezuelan migrants $<$5 years	1,122	0.843	1.335	0.000	10.660
Panel C: Election outcomes in 2018 second-round					
Votes for right	1,122	0.610	0.198	0.046	0.944
Votes for left	1,122	0.346	0.201	0.029	0.941
Turnout	1,122	0.520	0.095	0.076	0.742
Panel D: Election outcomes in 2018 first-round					
Votes for Duque	1,122	0.469	0.185	0.000	0.871
Votes for Petro	1,122	0.237	0.184	0.011	0.879
Votes for Fajardo	1,122	0.130	0.095	0.000	0.492
Votes for de la Calle	1,122	0.020	0.020	0.000	0.276
Votes for Vargas	1,122	0.101	0.089	0.000	0.633
Votes for Trujillo	1,122	0.003	0.002	0.000	0.034
Panel E: Election outcomes in 2014 first-round					
Votes for right	1,122	0.337	0.164	0.000	0.812
Votes for center	1,122	0.498	0.158	0.000	1.000
Votes for left	1,122	0.097	0.068	0.000	0.616
Panel F: Variation with previous elections					
Var. votes CD (1st round)	1,122	0.132	0.099	-0.228	0.499
Var. votes peace plebiscite	1,122	0.009	0.128	-0.424	0.506

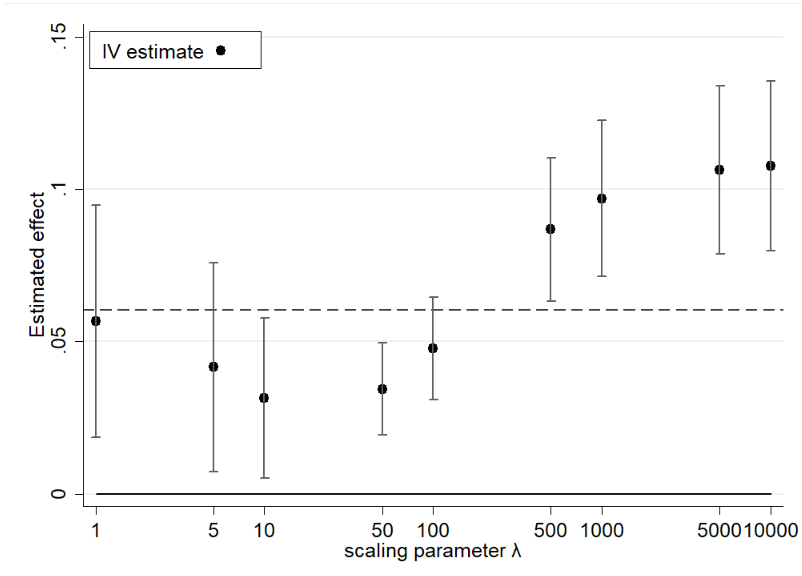
Notes: The table present the summary statistics for the main control variables by its classification. The observations are at the municipal level.

where λ is a scaling parameter. Note, when $\lambda = 1$ the decay function takes values close to zero as the distance increases. In particular, the $distance$ variable takes values between 3 and 1135 with mean equal to 414.6; this translate into a decay function with mean equal 0.0000271 and values ranging from 0 to 0.0304. Given the lower values, I use different values for the scaling parameters λ and show that the results are not driven by the value chosen for it. First-stage estimates are robust to the value of the scaling parameter, although all of them present lower values than that of my main specification; while second-stage estimates are more sensible to this value, yet there are around the estimated parameter from my main specification.

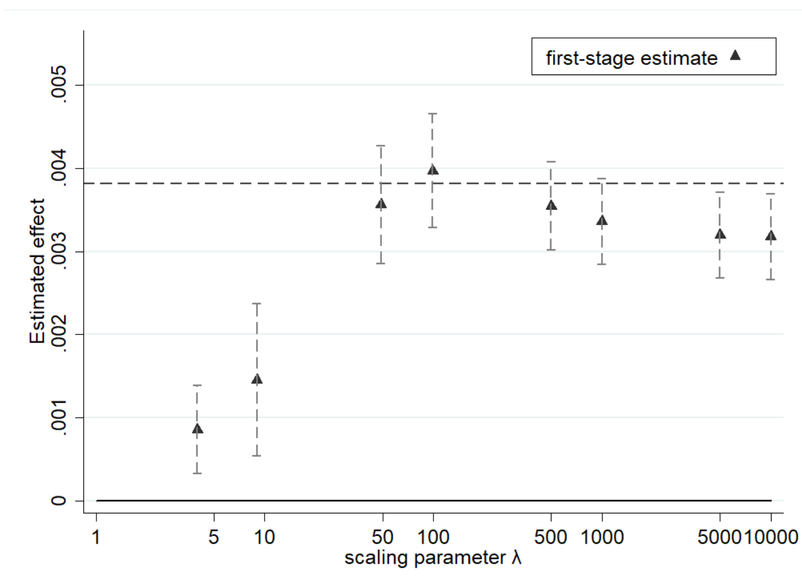
Second round results – Table B.3 presents the IV estimates of equation 3.1 using the election results of the second round of the 2018 presidential election, along with the OLS estimates, the reduce form, and the first-stage from equation 3.2. The two candidates are classified as in the first-round following Fergusson et al. (2021) criteria. Ivan Duque is a right-wing candidate and Gustavo Petro is a left-wing candidate, therefore the results at the candidate level are the same as those for the political preferences. The results mimic those of the first round, i.e. presence of Venezuelans shift votes from the left-wing candidate towards the right-wing candidate while the turnout increases.

Figure B.1: Effects of Venezuelan presence using different exponential decay functions

(a) Second-stage



(b) First-stage



Note: Each point represents a different estimation. The horizontal axis report the value of taken by the scaling parameter λ. Horizontal dash lines represent the estimated values of the main specification.

Table B.3: Effects of presence of Venezuelan migrants on presidential election (second-round)

Dependent variable:	Votes for right		Votes for left		Total votes	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2SLS IV regression (second-stage)						
Presence Venezuelan migrants	0.142***	0.059***	-0.143***	-0.059***	0.051***	0.037***
	(0.016)	(0.009)	(0.016)	(0.009)	(0.009)	(0.008)
Panel B: OLS regression						
Presence Venezuelan migrants	0.034***	0.023***	-0.032***	-0.021***	0.020***	0.019***
	(0.008)	(0.006)	(0.007)	(0.005)	(0.004)	(0.004)
Panel C: Reduced form						
Distance function	0.064***	0.028***	-0.064***	-0.027***	0.023***	0.017***
	(0.007)	(0.005)	(0.007)	(0.005)	(0.004)	(0.004)
Panel D: 2SLS first-stage regression						
Distance function	0.449***	0.467***	0.449***	0.467***	0.449***	0.467***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
F-test (excluded instruments)	139.88	151.87	139.88	151.87	139.88	151.87
Observations	1,035	1,035	1,035	1,035	1,035	1,035
Peace plebiscite		✓		✓		✓
Border controls	✓	✓	✓	✓	✓	✓
Distance controls	✓	✓	✓	✓	✓	✓
Administrative controls	✓	✓	✓	✓	✓	✓
Mayor election controls	✓	✓	✓	✓	✓	✓

Notes: Each entry is estimated separately. Panel A and B present estimates of equation 3.1 and panel D of equation 3.2. Dependent variable is reported in the column header. First-stage estimates uses an inverse logarithmic decay function of the distance to the closest entry port. All columns include all the control variables but odd columns do not control for the results of the peace plebiscite. Robust standard errors presented in parenthesis. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Bibliography

Abadie, Alberto, Maria C. Acevedo, Maurice Kugler, and Juan F. Vargas. 2015. "Inside the War on Drugs: Effectiveness and Unintended Consequences of a Large Illicit Crops Eradication Program in Colombia." Working Paper Harvard Kennedy School.

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2022. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics*, 138(1): 1–35.

Acevedo, Karina Margarita, and Iván David Bornacelly. 2014. "Panel Municipal del CEDE." Universidad de los Andes - CEDE Serie Documentos CEDE 2014-26.

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019. "Shift-Share Designs: Theory and Inference." *The Quarterly Journal of Economics*, 134(4): 1949–2010.

Aldeco, Lorenzo, Jose A. Jurado, and Aurora A. Ramírez-Álvarez. 2022. "Internal Migration and Drug Violence in Mexico." Banco de México Working Papers 2022-11.

Althaus, Dudley, and Steven Dudley. 2014. "Mexico's Security Dilemma: The Rise of Michoacan's Militias." InSightCrime.

Angrist, Joshua D., and Adriana D. Kugler. 2008. "Rural Windfall or a New Resource Curse? Coca, Income, and Civil Conflict in Colombia." *The Review of Economics and Statistics*, 90(2): 191–215.

APHIS. 2016. "Mexican Hass Avocado Import Program, Docket No. APHIS-2014-0088."

Asfura-Heim, Patricio, and Ralph Espach. 2013. "The Rise of Mexico's Self-Defense Forces: Vigilante Justice South of the Border." *Foreign Affairs*, 92(4): 143–150.

Atuesta, Laura H., and Aldo F. Ponce. 2017. "Meet the *Narco*: increased competition among criminal organisations and the explosion of violence in Mexico." *Global Crime*, 18(4): 375–402.

Atuesta, Laura H., and Yocelyn Samantha Pérez-Dávila. 2017. "Fragmentation and cooperation: the evolution of organized crime in Mexico." *Trends in Organized Crime*, 21(3): 235–261.

Atuesta, Laura H., Oscar S. Siordia, and Alejandro Madrazo Lajous. 2018. "The "War on Drugs" in Mexico: (Official) Database of Events between December 2006 and November 2011." *Journal of Conflict Resolution*, 63(7): 1765–1789.

Bahar, Dany, Ana María Ibáñez, and Sandra V. Rozo. 2021. "Give me your tired and your poor: Impact of a large-scale amnesty program for undocumented refugees." *Journal of Development Economics*, 151: 102652.

Barone, Guglielmo, Alessio D'Ignazio, Guido de Blasio, and Paolo Naticchioni. 2016. "Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior." *Journal of Public Economics*, 136: 1 – 13.

Battiston, Giacomo, Gianmarco Daniele, Marco Le Moglie, and Paolo Pinotti. 2022. "Fueling Organized Crime: The Mexican War on Drugs and Oil Thefts." *SSRN Electronic Journal*.

BBC. 2018. "Venezuela: 3 gráficos que muestran la enorme dimensión del éxodo en los últimos años por culpa de la crisis." BBC Mundo. August 28th, 2018. Accessed from: https://www.bbc.com/mundo/noticias-america-latina-45262411.

BBC. 2019. "Venezuela crisis: Four million have fled the country, UN says." BBC News. June 7th, 2019. Accesed from: https://www.bbc.com/news/world-latin-america-48559739.

Bello, Omar D., Juan S. Blyde, and Diego Restuccia. 2011. "Venezuela's Growth Experience." *Latin American Journal of Economics*, 48(2): 199–226.

Bernal, Carolina, Mounu Prem, Juan F. Vargas, and Monica Lorena Ortiz Medina. 2022. "Peaceful Entry: Entrepreneurship Dynamics during Colombia's Peace Agreement." *SSRN Electronic Journal*.

Blattman, Christopher. 2022. *Why We Fight The Roots of War and the Path to Peace*. Penguin Books, Limited.

Boone, Christopher, Heidi Kristiina Kaila, and David E. Sahn. 2022. "Posh spice or scary spice? The impacts of Madagascar's vanilla boom on household well-being." *SSRN Electronic Journal*.

Borusyak, Kirill, and Peter Hull. 2023. "Nonrandom Exposure to Exogenous Shocks." *Econometrica*, 91(6): 2155–2185.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2021. "Quasi-Experimental Shift-Share Research Designs." *The Review of Economic Studies*, 89(1): 181–213.

Buonanno, Paolo, Ruben Durante, Giovanni Prarolo, and Paolo Vanin. 2015. "Poor Institutions, Rich Mines: Resource Curse in the Origins of the Sicilian Mafia." *The Economic Journal*, 125(586): F175–F202.

Calderón, Gabriela, Gustavo Robles, Alberto Díaz-Cayeros, and Beatriz Magaloni. 2015. "The Beheading of Criminal Organizations and the Dynamics of Violence in Mexico." *Journal of Conflict Resolution*, 59(8): 1455–1485.

Camacho, Adriana, and Daniel Mejía. 2017. "The health consequences of aerial spraying illicit crops: The case of Colombia." *Journal of Health Economics*, 54: 147 – 160.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64.

Castillo, Tomás, and Mercedes Reguant. 2017. "Percepciones sobre la migración venezolana: causas, España como destino, expectativas de retorno." *Migraciones. Publicación del Instituto Universitario de Estudios sobre Migraciones*, , (41): 133–163.

Chabat, Jorge. 2010. "La respuesta del gobierno de Calderón al desafío del narcotráfico: entre lo malo y lo peor." Centro de Investigación y Docencia Económicas, División de Estudios Internacionales 196.

Charlton, Diane, Alexander James, and Brock Smith. 2021. "Seasonal agricultural activity and crime." *American Journal of Agricultural Economics*, 104(2): 530–549.

Cisneros, Elías, Tobias Hellmundt, and Krisztina Kis-Katos. 2022. "Conflicts of rural transformation: The effects of oil palm expansion in Indonesia."

CNN. 2019. "Ya son más de 4 millones de migrantes y refugiados venezolanos en el mundo, según ACNUR." CNN en Español. June 7th, 2019. Accessed from: https://cnnespanol.cnn.com/2019/06/07/mas-de-4-millones-de-migrantes-y-refugiados-venezolanos-en-el-mundo-segun-acnur/.

Conley, T.G. 1999. "GMM estimation with cross sectional dependence." *Journal of Econometrics*, 92(1): 1–45.

Coscia, Michele, and Viridiana Rios. 2012. "Knowing where and how criminal organizations operate using web content." 1412–1421. ACM Press.

Cote, Juan Pablo. 2019. "The Effect of Interdiction on Coca Cultivation in Colombia." Universidad de los Andes - CEDE Serie Documentos CEDE 2019-18.

Crost, Benjamin, and Joseph H. Felter. 2019. "Export Crops and Civil Conflict." *Journal of the European Economic Association*, 18(3): 1484–1520.

Dammert, Ana C. 2008. "Child labor and schooling response to changes in coca production in rural Peru." *Journal of Development Economics*, 86(1): 164 – 180.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2018. "Fuzzy Differences-in-Differences." *The Review of Economic Studies*, 85(2): 999–1028.

de Chaisemartin, Clément, Xavier D'Haultfœuille, and Yannick Guyonvarch. 2019. "Fuzzy differences-in-differences with Stata." *The Stata Journal*, 19(2): 435–458.

De Haro, Itzel. 2022. "Avocados: Mexico's green gold.The U.S. opioid crisis and its impact on Mexico's drug cartel."

Dell, Melissa. 2015. "Trafficking Networks and the Mexican Drug War." *American Economic Review*, 105(6): 1738–1779.

Dimico, Arcangelo, Alessia Isopi, and Ola Olsson. 2017. "Origins of the Sicilian Mafia: The Market for Lemons." *The Journal of Economic History*, 77(4): 1083–1115.

DIRAN. 2014. *Coca: deforestación, contaminación y pobreza : acercamiento a la actividad agronómica y la problemática ambiental de los cultivos de coca en Colombia.* Policía Nacional, Dirección de Antinarcóticos.

Dixit, Avinash K. 2004. *Lawlessness and Economics.* Princeton University Press.

Dube, Oeindrila, and Juan Fernando Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." *The Review of Economic Studies*, 80(4): 1384–1421.

Dube, Oeindrila, Omar García-Ponce, and Kevin Thom. 2016. "From Maize to Haze: Agricultural Shocks and the Growth of the Mexican Drug Sector." *Journal of the European Economic Association*, 14(5): 1181–1224.

Dustmann, Christian, Kristine Vasiljeva, and Anna Piil Damm. 2018. "Refugee Migration and Electoral Outcomes." *The Review of Economic Studies*, 86(5): 2035–2091.

Dávalos, Eleonora, and Liliana M. Dávalos. 2020. "Social Investment and Smallholder Coca Cultivation in Colombia." *The Journal of Development Studies*, 56(6): 1118–1140.

Edo, Anthony, Yvonne Giesing, Jonathan Öztunc, and Panu Poutvaara. 2019. "Immigration and electoral support for the far-left and the far-right." *European Economic Review*, 115: 99–143.

Eggers, Andrew C., Guadalupe Tuñón, and Allan Dafoe. 2023. "Placebo Tests for Causal Inference." *American Journal of Political Science*.

Espinosa, Nicolás. 2010. *Política de vida y muerte: etnografía de la violencia diaria en la Sierra de la Macarena.* Instituto Colombiano de Antropología e Historia.

Estancona, Chelsea, and Lucía Tiscornia. 2022. "From Cocaine to Avocados: Criminal Market Expansion and Violence."

Fergusson, Leopoldo, Pablo Querubin, Nelson A. Ruiz, and Juan F. Vargas. 2021. "The Real Winner's Curse." *American Journal of Political Science*, 65(1): 52–68.

Ferraz, Eduardo, Rodrigo R. Soares, and Juan F. Vargas. 2022. "Unbundling the Relationship between Economic Shocks and Crime." *SSRN Electronic Journal*.

Ferri, Benjamin. 2022. "Novel Shift-Share Instruments and Their Applications."

Forero, Sebastián. 2019. "Escuelas en ruinas y rodeadas de coca, así estudian en la cordillera de Nariño." El Espectador. 06/25/2019. Available from: https://www.elespectador.com/colombia2020/territorio/escuelas-en-ruinas-y-rodeadas-de-coca-asi-estudian-en-la-cordillera-de-narino-articulo-867597 (last accesed on 06/28/2019).

Gallego, Jorge, Juan D. Martínez, Kevin Munger, and Mateo Vásquez-Cortés. 2019. "Tweeting for peace: Experimental evidence from the 2016 Colombian Plebiscite." *Electoral Studies*, 62: 102072.

Gelbach, Jonah B. 2016. "When Do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics*, 34(2): 509–543.

Gerdes, Christer, and Eskil Wadensjö. 2010. "The impact of immigration on election outcomes in Danish municipalities." Stockholm University, Linnaeus Center for Integration Studies - SULCIS SULCIS Working Papers 2010:3.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review*, 110(8): 2586–2624.

Guardado, Jenny, and Steven Pennings. 2020. "The Seasonality of Conflict." World Bank, Washington, DC.

Halla, Martin, Alexander F. Wagner, and Josef Zweimüller. 2017. "Immigration and Voting for the Far Right." *Journal of the European Economic Association*, 15(6): 1341–1385.

Harari, Mariaflavia, and Eliana La Ferrara. 2018. "Conflict, Climate, and Cells: A Disaggregated Analysis." *The Review of Economics and Statistics*, 100(4): 594–608.

Hastings, Justin, and David Ubilava. 2023. "Agricultural Shocks and Social Conflict in Southeast Asia."

Hirshleifer, Jack. 1991. "The Technology of Conflict as an Economic Activity." *The American Economic Review*, 81(2): 130–134.

Hirshleifer, Jack. 2001. *The Dark Side of the Force.* Cambridge University Press.

Hsieh, Chang-Tai, Edward Miguel, Daniel Ortega, and Francisco Rodriguez. 2011. "The Price of Political Opposition: Evidence from Venezuela's "Maisanta"." *American Economic Journal: Applied Economics*, 3(2): 196–214.

Huang, Kuan-Ming, Zhengfei Guan, Trent Blare, and A. Malek Hammami. 2023. "Global Avocado Boom." *Choices*, 38(4): 1–9.

Ibanez, Marcela, and Fredrik Carlsson. 2010. "A survey-based choice experiment on coca cultivation." *Journal of Development Economics*, 93(2): 249 – 263.

Ibáñez, Marcela. 2010. "Who crops coca and why? The case of Colombian farmers." Courant Research Centre: Poverty, Equity and Growth Courant Research Centre: Poverty, Equity and Growth - Discussion Papers 40, Göettingen.

Isacson, Adam. 2019. "Restarting Aerial Fumigation of Drug Crops in Colombia is a Mistake." Washington Office on Latin America WOLA. 03/7/2019. Available from: https://www.wola.org/analysis/restarting-aerial-fumigation-of-drug-crops-in-colombia-is-a-mistake/ (last accesed on 07/26/2019).

Jaeger, David, Joakim Ruist, and Jan Stuhler. 2018. "Shift-Share Instruments and the Impact of Immigration." National Bureau of Economic Research.

Knight, Brian, and Ana Tribin. 2023. "Immigration and violent crime: Evidence from the Colombia-Venezuela Border." *Journal of Development Economics*, 162: 103039.

Kruger, Diana I. 2007. "Coffee production effects on child labor and schooling in rural Brazil." *Journal of Development Economics*, 82(2): 448–463.

Lebow, Jeremy. 2022. "The labor market effects of Venezuelan migration to Colombia: reconciling conflicting results." *IZA Journal of Development and Migration*, 13(1).

Lebow, Jeremy, Jonathan Moreno Medina, and Horacio Coral. 2020. "Immigration and Trust: The Case of Venezuelans in Colombia." *SSRN Electronic Journal*.

Lessing, Benjamin. 2015. "Logics of Violence in Criminal War." *Journal of Conflict Resolution*, 59(8): 1486–1516.

Lessing, Benjamin. 2020. "Conceptualizing Criminal Governance." *Perspectives on Politics*, 19(3): 854–873.

López, Mónica, Carlos Andrés Guarín, Oscar Iván Medina, and Hernando Zuleta. 2019. "Proceso de paz con las FARC y cultivos de coca en el período 2013-2017." Universidad de los Andes - CEDE Serie Documentos CEDE 2019-32.

Mayda, Anna Maria, Giovanni Peri, and Walter Steingress. 2022. "The Political Impact of Immigration: Evidence from the United States." *American Economic Journal: Applied Economics*, 14(1): 358–389.

McKenzie, David. 2023. *Interpreting treatment effects on an inverse hyperbolic sine outcome variable and alternatives.* Last accessed: September 1, 2023. https://blogs.worldbank.org/impactevaluations/interpreting-treatment-effects-inverse-hyperbolic-sine-outcome-variable-and.

Mehlum, Halvor, Karl Moene, and Ragnar Torvik. 2006. "Institutions and the Resource Curse." *The Economic Journal*, 116(508): 1–20.

Mejía, Daniel, and Carlos Esteban Posada. 2008. "Cocaine Production and Trafficking: What Do We Know?" World Bank Policy Research Working Paper Series 4618. Available at SSRN: https://ssrn.com/abstract=1149121.

Mejía, Daniel, and Pascual Restrepo. 2015. "Bushes and Bullets: Illegal Cocaine Markets and Violence in Colombia."

Mejía, Daniel, Pascual Restrepo, and Sandra V. Rozo. 2017. "On the Effects of Enforcement on Illegal Markets: Evidence from a Quasi-Experiment in Colombia." *World Bank Economic Review*, 31(2): 570–594.

MEN. 2013. "Sistema nacional de indicadores educativos para los niveles de preescolar, básica y media en Colombia." Ministerio de Educación Nacional.

MEN. n.d.. "La deserción escolar." Ministerio de Educación Nacional, unpublished document. Available from: http://www.mineducacion.gov.co/1621/articles-293659_archivo_pdf_abc.pdf.

Mendez, Ildefonso, and Isabel M. Cutillas. 2014. "Has immigration affected Spanish presidential elections results?" *Journal of Population Economics*, 27(1): 135–171.

Miller, Jeff. 2020. *Avocado: A Global History. The Edible Series*, Reaktion Books, Limited.

Millimet, Daniel L., and Marc F. Bellemare. 2023. "Fixed Effects and Causal Inference."

OECD. 2016. "Education in Colombia." OECD Publishing.

OLR. 2017. "Características de los migrantes de Venezuela a Colombia." Observatorio Labroral de la Universidad del Rosario techreport 3.

Olson, Mancur. 1993. "Dictatorship, Democracy, and Development." *American Political Science Review*, 87(3): 567–576.

Otto, Alkis Henri, and Max Friedrich Steinhardt. 2014. "Immigration and election outcomes - Evidence from city districts in Hamburg." *Regional Science and Urban Economics*, 45: 67 – 79.

Piccato, Pablo. 2022. *Historia mínima de la violencia en México.* El Colegio de México.

Prem, Mounu, Juan F. Vargas, and Daniel Mejía. 2023. "The Rise and Persistence of Illegal Crops: Evidence from a Naive Policy Announcement." *The Review of Economics and Statistics*, 105(2): 344–358.

Prem, Mounu, Juan F. Vargas, and Olga Namen. 2023. "The Human Capital Peace Dividend." *Journal of Human Resources*, 58(3): 962–1002.

Prieto-Curiel, Rafael, Gian Maria Campedelli, and Alejandro Hope. 2023. "Reducing cartel recruitment is the only way to lower violence in Mexico." *Science*, 381: 1312–1316.

Ratha, Dilip, and William Shaw. 2007. *South-South Migration and Remittances.* The World Bank.

Reyes, Luis Carlos. 2014. "Estimating the Causal Effect of Forced Eradication on Coca Cultivation in Colombian Municipalities." *World Development*, 61: 70 – 84.

Rios, Viridiana. 2013. "Why did Mexico become so violent? A self-reinforcing violent equilibrium caused by competition and enforcement." *Trends in Organized Crime*, 16(2): 138–155.

Rodríguez, Catherine, and Fabio Sánchez. 2010. "Books and Guns: The Quality of Schools in Conflict Zones." Universidad de los Andes Serie Documentos CEDE 2010-38.

Rodríguez, Claudia J. 2020. "The Effects of Aerial Spraying of Coca Crops on Child Labor, School Attendance, and Educational Lag in Colombia, 2008-2012." *Journal on Education in Emergencies*, 6(1): 84.

Rozo, Sandra, and Juan F. Vargas. 2020. "The Effect of Venezuelan Migration on Educational Outcomes in Colombia." Inter-American Development Bank.

Rozo, Sandra V., and Juan F. Vargas. 2021. "Brothers or invaders? How crisis-driven migrants shape voting behavior." *Journal of Development Economics*, 150: 102636.

SAGARPA. 2017. "Agucate Mexicano."

Salcedo, Dario. 2021. "The Unintended Consequences of Fighting Crime: Mexico's War on Drugs."

Santamaria, Julieth. 2020. "'When a Stranger Shall Sojourn with Thee': The Impact of theVenezuelan Exodus on Colombian Labor Markets." Job Market Paper.

Santos, Rafael J. 2018. "Blessing and curse. The gold boom and local development in Colombia." *World Development*, 106: 337–355.

Shirk, David, and Joel Wallman. 2015. "Understanding Mexico's Drug Violence." *Journal of Conflict Resolution*, 59(8): 1348–1376.

SIAP. 2018. "Atlas Agroalimentario 2012-2018." Servicio de Informacion Agroalimentaria y Pesquara (SIAP).

Sobrino, Fernanda. 2020. "Mexican Cartl Wars: Fighting for the U.S. Opioid Market."

Sviatschi, Maria Micaela. 2022. "Making a NARCO: Childhood Exposure to Illegal Labor Markets and Criminal Life Paths." *Econometrica*, 90(4): 1835–1878.

Ubilava, David, Justin V. Hastings, and Kadir Atalay. 2023. "Agricultural windfalls and the seasonality of political violence in Africa." *American Journal of Agricultural Economics*, 105(5): 1309–1332.

UNODC. 2014. "Colombia: Coca cultivation survey 2013." United Nations Office on Drugs and Crime, Bogota.

UNODC. 2017. "Colombia: Survey of territories affected by illicit crops - 2016." United Nations Office on Drugs and Crime, Bogota.

UNODC. 2019. "Colombia: Monitoreo de territorios afectados por cultivos ilícitos 2018." United Nations Office on Drugs and Crime, Bogotá.

Uribe-Castro, Mateo. 2021. "Caffeinated Development: Exports, Human Capital,and Structural Transformation in Colombia."

USDA. 2020. "Avocado Annual, Report Number: MX2020-0049."

Utar, Hale. 2018. "Firms and Labor in Times of Violence: Evidence from the Mexican Drug War." *SSRN Electronic Journal*.

Vargas, Juan F., Luis Fernando Gamboa, and Viviana García. 2014. "El lado oscuro de la equidad: violencia y equidad en el desempeño escolar." *Revista Desarrollo y Sociedad*, 74: 309–334.

Wainwright, Tom. 2016. *Narconomics.* Ebury Publishing.

Woldemikael, Olivia. 2022. "Migration as Municipal Management: The Political Effects of Venezuelan Migration in Colombia."

Wolff, Michael J. 2020. "Insurgent Vigilantism and Drug War in Mexico." *Journal of Politics in Latin America*, 12(1): 32–52.

Wu, Feng, Berdikul Qushim, Marcelo Calle, and Zhengfei Guan. 2018. "Government Support in Mexican Agriculture." *Choices*, 33(3): 1–11.

Zuleta, Hernando. 2019. "Coca, Cocaine and Drug Trafficking." Universidad de los Andes - CEDE Serie Documentos CEDE 2019-9.