Title

Three Essays on Agricultural Land and Labor Markets

Author

Arteaga, Julian

Publication Date

2024

Peer reviewed|Thesis/dissertation

Three Essays on Agricultural Land and Labor Markets

By

JULIAN ARTEAGA
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Agricultural and Resource Economics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Michael Carter, Chair

Ashish Shenoy, Chair

Stephen Boucher

Committee in Charge

2024

To my parents, Juan y María*, with love.*

Contents

Abstract

This dissertation is composed of three essays that study aspects related to the way in which agricultural land and labor markets operate, with a special emphasis on the particularities of agricultural input markets in developing countries. The first chapter investigates how government restrictions on land markets impact the agricultural sector, and assess whether such restrictions can curb distortions that stem from the presence of market power. To do so, I develop a general-equilibrium production model in which large landholders exert market power in both land and labor markets, and where there are limits on land accumulation. Restrictions reduce the inefficiencies arising from market power, but also hinder productive reallocation, with the net effect on productivity depending on initial levels of land concentration. I empirically test the model's predictions by estimating how a law imposing municipality-specific limits on landholdings in Colombia affected productivity, land concentration, and agricultural labor markets. To estimate the impact of the law, I combine a collection of rich micro-level data sources which include a newly built dataset on municipal agricultural productivity. Exploiting plausibly exogenous variation in restriction stringency across bordering municipalities, I find that imposing restrictions caused a permanent reduction in productivity and only modest reductions in overall land inequality. However, restrictions also increased both agricultural workers' earnings and the employment share in agriculture, suggesting they were beneficial to landless wage laborers by reducing labor market power.

The second chapter (co-authored with Nicolás de Roux, Margarita Gáfaro, Ana María Ibáñez, and Heitor Pellegrina) studies the effect of weather shocks on rural land sales and the farm size distribution. Using a unique administrative dataset with transaction-level information and a land registry covering most of Colombia's farmland, it shows that extreme temperature events increase the frequency of land sales and decrease the average farm size within municipalities. These results are driven by small farms being subdivided and purchased by previously landless owners, with no evidence of weather shocks leading to the consolidation of small farms into larger holdings. The effects of extreme temperature on land sales are stronger in poorer and more isolated municipalities, where landowners are also less likely to take out land mortgages after a shock. To explain these patterns and explore how they can be exacerbated by underdevelopment, this chapter further develops an intertemporal, two-sector model where agents face a subsistence consumption constraint. Taken together, the model and the empirical findings preesented in this chapter highlight how climate-induced distress land sales are a relevant margin of adjustment that can have large distributional and efficiency implications for the agricultural sector of developing economies.

The third chapter (co-authored with Ashish Shenoy) focuses on the effect that variations in migrant worker inflows have on agricultural labor markets in destination economies. Using

information on migratory flows for every Mexican municipality and U.S. county pair throughout the 2006–2019 period, the chapter shows estimates on the effect that variations in Mexican migration flows have on U.S. agricultural labor-market outcomes. We instrument for migration-driven changes in local labor supply using a shift-share variable that combines Mexican municipality-level violence levels with preexisting migration network patterns. Our estimates show that, in the short run, decreasing migration rates put upward pressure on wages across all types of agricultural workers, and cause a large increase in the number of H-2A seasonal worker visas requested by employers. Conversely, in the long run, decreasing migration rates lead to lower wages in agriculture accompanied by slight reductions in employment levels. Regarding the mechanisms driving this result, we find that an exogenous decrease in the cumulative number of migrant arriving to a county during this period led to reductions in the acreage planted with labor-intensive crops, higher rates of mechanization, and lower average farmland values.

Acknowledgements

Not too long ago I came to the realization that my time as a Ph.D. student has been fundamentally happy. In large part this is due to the fact that, throughout my career, I've had the great fortune of being surrounded by exceptional mentors, colleagues, and friends. It hopefully does not come across as disingenuous when I say that it is because of them that this has been an interesting, worthy journey. I do not take this for granted, and so I would like to thank some people.

I've had the fortune of having Michael Carter and Ashish Shenoy as advisors. I owe a great debt of gratitude to Michael, who with an enormous amount of patience and generosity made me part of several research projects and showed me how independent research is done. Michael's intuition, depth, and impressive work capacity is something I have benefitted from greatly. Countless times I've been struck by his ability to take any idea and pass it through a lens that reveals much more economic nuance than what I had first thought. Most of all, Michael has left in me a sense that the work one does should matter, and that this is a responsibility to be taken seriously. It is in this idea where my deepest admiration for him lies.

I have not learned more economics from anyone else in my career than from Ashish. His capacity to break down any problem into distinct, straightforward components, and the naturality with which he can find the economic insight——or lack of it——in any idea, is something that has inspired awe in me since we first met. I have now become used to the experience of understanding a comment or a question he spontaneously makes about some issue only weeks later after having thought about it carefully. While his talent as an economist is evident to anyone who meets him, what I admire the most about Ashish is the deep, understated care he has for students and those who work with him. I have many things to thank Ashish for, and I feel honored to have been able to collaborate with him these years, something I hope will continue in the future.

Having worked and learned from Ana María Ibáñez since my time as an undergrad has been easily the single most important strike of professional luck I've ever had. It is because of her that I even began to consider the idea of pursuing a Ph.D., and ever since she has been a constant source of support and encouragement. The number of things I have to thank her for would fill many pages, and I will always be extremely proud of being part of the group of people she has mentored.

I would also like to thank many other UC Davis faculty members for their teachings and advice. Steve Boucher in particular has become to me one of the principal examples of how kindness and the utmost care for students is not only compatible, but also instrumental for successful scholarship. Tim Beatty, Andrés Carvajal, Dalia Ghanem, Travis Lybbert, Diana Moreira, Giovanni Peri, Aaron Smith, and Arman Rezaee all taught me many important things. I'm also immensely grateful to my coauthors, Nicolás de Roux, Margarita Gáfaro, Andrew Hobbs, Heitor Pellegrina, and Kelvin Shikuku. I hope to keep learning from all of you. Arnon Erba, Gabriela Garzón, Watson Lepariyo, Paola Poveda, and Laurie Warren also have all my gratitude.

In particular, my time in Davis was happy because I shared this journey with a group of people I will have in my heart forever. Aya and Gabi in Casa Colombia, which became home; Sarah, Dasha, and Sangwon always there from day one; Charlotte and Ethan who adopted me. I am truly grateful to Shisham Adhikari, Isaac Ahimbisibwe, Laura Alcocer, Sebastian Amador, Ellen Anderson, TJ Anderson, Inés Badin, Oscar Barriga, Beau Bressler, Carlos Brito, Alessandro Caiumi, Kalyani Chaudhuri, Saloni Chopra, Diógenes Cruz, Miki Doan, Enrique Estefanía, Mari Fenton, Felipe Grimaldi, James Hasbany, Jeff Hadachek, Giuseppe Ippedico, Colter Hoffman, Rachel Jones, Caitlin Kieran, Mira Korb, Tina Kotsakou, Pedro Liedo, Marcos Martínez, Laura Meinzen-Dick, Sara Meza, Aleks Michuda, Shotaro Nakamura, Karen Ortiz-Becerra, Nacho Oliva, Anderson Ospino, Stephen Paolillo, Socorro Pardo, Isy Picciotto, Bruno Pimenta, Fran Pizzi, Joe Raymond, Nicolás Restrepo, Devin Serfas, Scott Sommerville, Reid Taylor, Will Troske, Madeline Turland, Joakim Weill, Kelly Wu, Reem Zaiour, Geyi Zheng, Yanan Zheng, and to so many more friends and colleagues—too many to name—, both from the ARE and Econ departments, from cohorts above and below, here in California and back in Bogotá. I would like to also dedicate this dissertation to the memory of Rafael Santos and Sebastian Ramirez Amaya, both of whom, in their own way, taught me so much.

My mother, my father, and my brother, María, Juan, and Martín, have given me the kind of support and motivation that can only be found in unconditional love. They are my foundation; all of my accomplishments are thanks to them and are for them.

Finally, having gone through this journey by Tatiana Hiller's side has been one of the great privileges of my life. Day after day she has been there, facing all fears, dreaming together. I trust no one's judgement more, and any challenge is simpler with her. In the deepest sense, she makes me a better human being. For that I am grateful beyond words.

Chapter 1

Land-Market Restrictions and Agricultural Productivity under Market Power

1.1 Introduction

Governments often impose restrictions on the sale of assets being granted as part of redistributive policies. This practice is particularly prevalent in the distribution of agricultural land, where restrictions take varied forms such as use-contingent property rights or area limits on ownership. Although these policies are often justified as necessary to prevent land reconcentration, a growing body of evidence shows that these constraints hinder productivity and labor mobility across sectors, slowing economic development (de Janvry et al., 2015; Gottlieb and Grobovšek, 2019; Adamopoulos and Restuccia, 2020).

Rapid land reconcentration may however indicate the presence of market failures in rural land and labor markets. In particular, large owners may seek to accumulate land and operate inefficiently large farms in order to exert market power and distort input prices (Conning, 2003). If governments are politically constrained to directly implement perfect competition by breaking up large estates, and if the distortions produced by market power are larger than those introduced by restrictions, the application of second-best policies that disallow some types of land sales could be warranted (Lipsey and Lancaster, 1956). While the practice of restricting land markets is prevalent and has been employed by many governments across the world (Allen, 1991), it is still uncertain if imperfect-competition arguments used as justification hold validity.

Is the presence of market power a large enough concern to merit the restriction of land transactions? In this paper I investigate this question by studying how the imposition of land market restrictions affects land and labor markets, land concentration, and agricultural productivity. I focus on estimating the impact of a 1994 law—Law 160—that established land ceilings of varying height and on varying amounts of farmland across hundreds of municipalities in Colombia, a country where both high land inequality levels and ambitious land distribution policies have been longstanding.

I first theoretically explore the importance of land market distortions by developing a general-equilibrium model where some agents exert market power to influence the price of land and labor, and where I introduce land market restrictions in the form of limits on individual landholdings within a fraction of the economy's farmland. Restrictions distort the efficient reallocation of inputs across producers, but also curtail market power. I show that the net effects of land market restrictions on productivity and agricultural wages are theoretically ambiguous and ultimately depend on the initial level of land concentration. At high enough concentration levels, the inefficiency caused by market power distortions exceeds that of misallocation in ownership, and the unrestricted equilibrium is less efficient than a restricted one.

To study this tradeoff empirically, I build a new dataset of detailed information on crop-specific agricultural yields in Colombia at the municipality-year level for the period 1988–2004. This dataset, built by digitizing and harmonizing hundreds of distinct archival government publications, constitutes the first comprehensive collection of municipal-level figures on agricultural output and area planted in the country for this time period. I further combine this data with information on the universe of public-land allocations made by the government throughout the twentieth century, as well as with several other sources of micro-level data on individual land sales, the distribution of farm sizes, and agricultural workers' earnings and employment.

I test the model's predictions by exploiting municipality-level variation in the stringency of market restrictions to estimate how Law 160 affected local rural economies in Colombia. Identification relies on two features of the law: First, ownership ceilings were only imposed on land originally owned by the state and subsequently publicly distributed. Thus municipality-level exposure varied with the fraction of land covered by the regulation. Second, the height of land ceilings was set to vary to account for differences in broadly defined regional agroclimatic conditions. The combination of these two features creates cross-sectional variation in the stringency of land market restrictions imposed on different municipalities.

To address the concern that restriction levels might be endogenous to unobserved municipal characteristics, the estimation restricts to municipality pairs that share a border. Under the assumption that endogenous characteristics vary smoothly across municipal borders, discrete differences in ceiling height and share of farmland under restriction between neighboring municipalities serve as a source of plausibly exogenous variation in the stringency of market constraints. My econometric strategy therefore follows a differences-in-differences approach that relies on the within-municipality-pair variation in the severity of restrictions to estimate how the law affected agricultural productivity, land inequality, and agricultural workers' earnings and employment. Following the model's predictions, I further test for heterogeneity by initial concentration levels, which correspond to the strength of the incentive to reconcentrate after land reform. Supporting the identifying assumption of the difference-in-difference approach, I find no evidence of differential pretrends by treatment intensity.

I find that more stringent restrictions led to only slight reductions in overall land inequality, but caused a permanent reduction in agricultural land productivity. During the ten years following the passing of the law, a municipality moving from the 25^{th} to the 75^{th} percentile of the restriction stringency measure would have seen a decrease of roughly 7% in its land ownership gini index (about a quarter of a standard deviation in the distribution of gini indices across municipalities), but it would have suffered a persistent 25% reduction in revenue per hectare. Evidence suggests agricultural revenue per unit of land fell due to declining productivity within specific crops rather than through changes in municipal crop composition. These results represent, to the best of my knowledge, the first quantitative evaluation of the effect of Law 160 on Colombia's agricultural sector.

Despite the drop in productivity, restrictions also led to a substantial increase in the earnings and employment of agricultural wage workers. Going from the 25^{th} to the 75^{th} percentile of restriction stringency would entail a 23% increase in the share of a municipality's workers occupied in agriculture, and to an increase of 68% in monthly earnings per agricultural worker. This divergence between productivity and labor earnings is indicative of large landholders inefficiently restricting labor demand to lower wages. Consistent with the model's predictions, the negative productivity effects depend on the degree of pre-reform land concentration present in the local economy, with more initially-concentrated municipalities having lower productivity reductions due to the law. These results are consistent with a model in which land concentration distorts agricultural input markets.

Policymakers should be aware of the potentially large distortions that market power can have on rural economies with high levels of land concentration. This paper shows that, a priori, restricting land markets has ambiguous effects on the economy, and that there are important distributive implications related to the imposition (or the lifting) of restrictions in contexts where imperfect competition might be prevalent. In particular for the Colombian context, the findings in this paper suggest that the policy of restricting land transfers has, on average, held back the efficiency of the agricultural sector, but that the rise in workers' earnings has benefitted landless wage laborers.

This paper contributes to the large literature on monopsony power in labor markets (e.g., Berger et al. (2022); Naidu et al. (2016), see Card (2022) for a review). There is growing evidence that firms exert labor market power in many settings, and even in markets that appear very competitive (Dube et al., 2020). In developing countries, and specifically in Latin America, Felix (2021), and Amodio and de Roux (2023) quantify the substantial degree of monopsony power held by firms. These estimates are based on mostly non-agricultural, urban jobs. I extend this type of analysis to rural workers, who are subject to large mobility frictions between, and even within, labor markets (Imbert and Papp, 2020; Emerick et al., 2022). Latin American agricultural workers—a region with some of the highest levels of land concentration in the world (Eslava and Caicedo, 2023)—potentially face labor markets where imperfect competition frictions are at least as prevalent as in urban areas. As a result, if market power distortions are large enough, an unrestricted land market could actually lead to an equilibrium that is both less equitable *and* less efficient (Deininger, 2003;

Carter and Zegarra, 2000).

This paper also contributes to the literature that investigates the causes of the large observed differences in agricultural productivity across countries (Restuccia et al., 2008; Gollin et al., 2014). Recent work has shown that this productivity gap can be in part explained by the imposition of land-tenure institutions that impede the allocation of resources towards activities with higher returns (de Janvry et al., 2015; Gottlieb and Grobovšek, 2019; Adamopoulos and Restuccia, 2020; Adamopoulos et al., 2022b; Chari et al., 2021; Adamopoulos et al., 2022a). This paper provides evidence on the productivity impacts of a previously unexplored policy effort aimed at restricting land markets. It also contributes to this literature by theoretically exploring how these policies interact with other sources of misallocation such as those stemming from market power in environments of high land concentration.

By showing that many of the observed positive effects of land reform on wages and poverty reduction are consistent with a model in which reforms curtail market power, this paper contributes to the literature that studies the effects of agrarian and property-rights reform in developing economies (see, for example, Besley and Burgess (2000); Banerjee et al. (2002); Carter and Olinto (2003); Deininger et al. (2008); Besley et al. (2016); Ortiz-Becerra (2021); Montero (2022)). Much of this literature studies the effect of reforms under the light of potential failures in credit markets, or as introducing changes in the relative bargaining power of tenants and landlords. I show that curbing distortions stemming from imperfect competition might be one additional channel through which land reforms can impact rural economies. Analyzing the effects of a closely-related but independent reform effort in Colombia, Galán (2018) estimates the micro-level intragenerational effects of increasing land access for rural workers. I contribute to this effort by estimating the joint general-equilibrium effects that land-market reform has on productivity, inequality, and agricultural workers' earnings.

Finally, this paper contributes to the study of the causes of persistence in land concentration and land utilization patterns across time. Assunção (2008b) and Bardhan et al. (2014) show how even very ambitious land reform efforts may not lead to persistent reductions in land inequality, while Smith (2020) shows how initial land endowments have lasting effects long after any restrictions on land sales have been lifted. Studying the Colombian case, Faguet et al. (2020) document how initial land concentration levels are a major determinant of the effectiveness of public-land allocation policies in improving land distribution measures, while Deininger (1999), and Assunção (2008a) further show that this concentration-persistence puzzle is compounded by the very stark underutilization of the land being concentrated. Following a conceptual framework that stresses the potential impacts of market power in rural economies with high land concentration levels—laid out in Conning (2003) and further empirically explored in Martinelli (2014)—this paper provides empirical evidence supporting the importance of this imperfect-competition mechanism in explaining these observed patterns.

1.2 Institutional Background

The unequal distribution of land in Colombia has historically been argued to be a major obstacle for economic development as well as one of the main drivers of violent conflict in the country; efforts to give poor farmers access to land have been numerous (Berry, 2017). Throughout the twentieth century the Colombian government tried in several occasions to reform landholding patterns using redistributive policies—either through direct state expropriation of large estates or by market-assisted reform—yet all of these attempts were largely unsuccessful (Ibáñez and Muñoz, 2010; CNMH, 2016).

1.2.1 Allocation of public land

By contrast, the free allocation of public idle lands (*baldíos*) to private individuals has been an uninterrupted policy of the Colombian state since the beginning of the twentieth century and has become, by far, the most consequential 'land reform' policy instrument employed by the government. This allocation process has mostly consisted of a combination of frontier-settlement schemes where unused public lands are granted to poor smallholders, and of programs focused on the titling of state-owned lands that might have been previously informally occupied (Albertus, 2015). Since the enactment of the Social Agrarian Reform Act (Law 135) in 1961, the explicit objective of the policy has been that of reducing land inequality and giving land to landless farmers.

Procedurally, land petitioners must go through an administrative process managed by the National Land Agency (ANT) that is meant to rule if petitioners fulfill the legal requirements to become the beneficiaries of an allocation. While the exact requirements have changed over time, petitioners have always been required to demonstrate they own no other land, and that they belong to a low-income household. Under the current legislation, the process formally consists of nine steps, which include the placement of an advertisement announcing the allocation in a local newspaper, and a physical inspection of the plot to be granted. Formally this procedure should take at most 60 days, but allocation processes are generally much lengthier and some can take years (Gutiérrez Sanín, 2019).

In terms of the number of beneficiaries and the amount of land allocated, the scale of the policy has been vast. Colombia has had "one of the Western Hemisphere's largest public land distribution programs during the last century" (Albertus, 2019), having granted (throughout the period 1901–2012) more than 500,000 land plots to private individuals in 1,031 of the 1,122 existing municipalities, amounting to roughly half of the currently privately-held land in the country (Sánchez and Villaveces, 2016; Arteaga et al., 2017). Despite its scale, the *baldío* allocation program did not fundamentally alter the country's starkly unequal land distribution, suggesting many of the plots allocated became reconcentrated with time (Faguet et al., 2020; Ibánez et al., 2012).

In contrast with other land distribution policies frequently imposed across the developing world (e.g. Adamopoulos et al. (2022a)), receiving and maintaining property rights over a *baldío* was not conditional on its direct use and cultivation, although recipients who sold their land were ineligible to receive any other state land for 15 years. Examining the effects of a parallel but independent land allocation program, Galán (2018) finds that ten years after receiving land, 30% of beneficiaries had indeed sold the plot to a third party.[1]

1.2.2 Land ownership ceilings - the Agricultural Family Unit (UAF)

Driven by the ineffectiveness of the public-land allocation policy to reduce land concentration during the previous three decades, the enactment of law 160 in 1994 established municipality-specific

[1]The *Sharecroppers and Tenants Program* studied by Galán (2018) did, in fact, impose a 10-year restriction on sales.

ceilings on the amount of land originally allocated by the government that any individual could, from that moment onwards, purchase and own. The ceiling was notionally defined as the amount of land that a rural household would require to obtain a minimum basic level of income and became known as the Agricultural Family Unit (UAF). It was established that the height of this ceiling would vary geographically to account for differences in agroecological conditions, and its magnitude was defined following the concept of 'relatively homogeneous zones', a novel geographical division that did not correspond to the traditional administrative divisions of *municipio* or *departamento*.[2]

Importantly, the law established that the ceiling on landholdings only applied to land that at some point in the past had been part of the public land distribution program. Land plots not initially allocated by the government were excluded from the restriction and no constraints were placed on how much of this type of land could be owned by individuals.

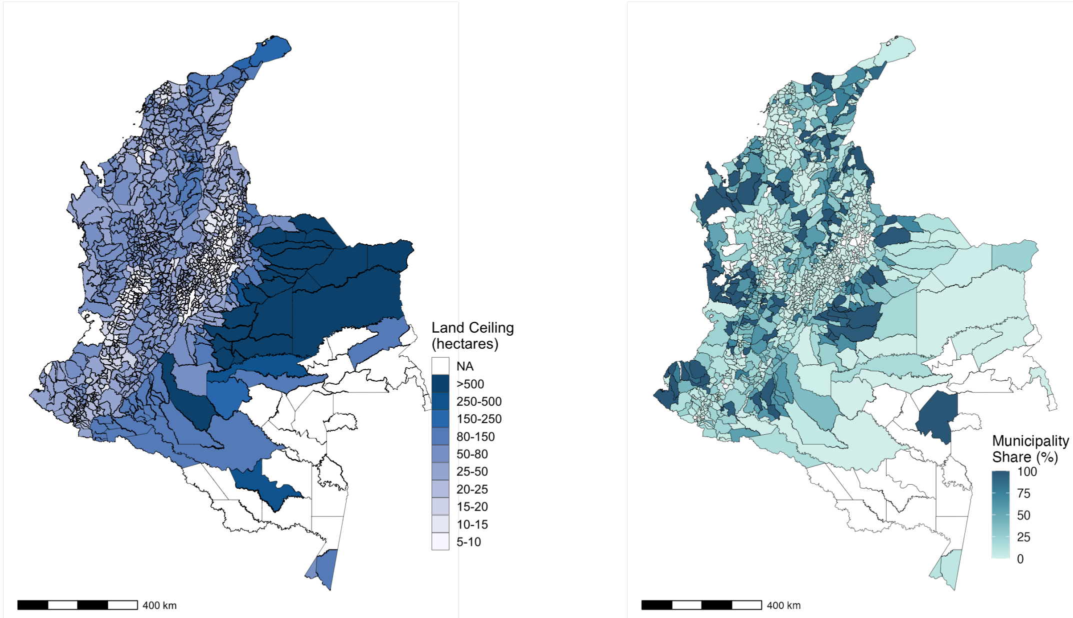
In practice, the restriction banned any future land transactions which would have resulted in an individual accumulating an amount of land above the stipulated ceiling. It did not, however, lead to any retroactive expropriation and redistribution of above-ceiling landholdings. At the same time, any transfer of land (either restricted or unrestricted) between individuals whose landholdings remained below the ceiling were also allowed.

The specific height of the land ceiling imposed in each municipality was formally published in resolution 041 of 1996 by the *Instituto Colombiano de la Reforma Agraria* (INCORA), the national land agency of the time. The amounts of land defined by these ceilings, as well as the restrictions on sales stipulated by law 160 are still currently in force. Legislation (law 902) passed in 2017 additionally established a complete ban on the sale of newly granted government plots for a period of seven years after the allocation has taken place. A detailed exposition of the institutional context in which law 160 was enacted can be found in chapter 3 of CNMH (2016). Figure 1.1 shows the geographical variation of both land-ceiling heights and of the share of land subject to restriction across Colombian municipalities.

Finally, the UAF restriction was only imposed on purchases, and no limits on land rental markets were imposed. Colombian land rental markets are, however, characterized by being notoriously thin,

[2]The second-level administrative unit in Colombia, *departamentos*, are equivalent to the U.S. states. *Departamentos* are composed of municipalities, which are analogous to U.S. counties.

Figure 1.1: Measures of Land-Market Restrictions at the Municipal Level.



(a) Ceiling Height (b) % of Government-Allocated Area in 1990

Notes. Geographical distribution of land-market restrictions. Left panel: Maximum UAF in hectares as defined by *INCORA* Resolution 041 of 1996. Right panel: Share of cumulative government-allocated land in 1990 as a fraction of total privately-owned farmland in the municipality according to the 2014 national agricultural census (CNA).

with only 9% of farmers surveyed in a 2019 nationally-representative survey reporting to operate any rented land.[3] For this reason, the model presented in the following section abstracts from rental markets and focuses solely on land sales.

1.3 Theoretical Framework

This section develops a general-equilibrium, agricultural-production model that aims to capture the main institutional features of the Colombian land ceiling policy outlined in Section 1.2, and to highlight the potential tension between the distinct sources of inefficiency that policies seeking to regulate land markets might be confronted by. In the spirit of Conning (2003), the model allows for some firms in the economy to exert market power in both land and labor markets and shows how this leads to the existence of multiple equilibria depending on initial endowments. I then extend the model to allow for the imposition of a ceiling on land ownership on a fraction of the economy's farmland and analyze how this impacts input demand decisions, prices, and aggregate productivity.

[3]National Agricultural Survey (ENA), carried out by the National Statistical Agency (DANE); 2019-1 bulletin.

1.3.1 Market power with no land ceiling

Setup: Consider an agrarian production economy with population N and a fixed amount of land L, where all farmers produce an homogeneous, exogenously priced good which is the numeraire. Farmers are endowed with non-tradeable ability levels for agricultural production s, and non-agricultural production s^{out}. Both abilities are drawn from a population distribution of skills $(s, s^{out}) \sim ln\mathcal{N}(\mu, \Sigma)$. Each farmer is also endowed with uniform amounts of labor n_i^0 and varying amounts of land l_i^0.

Suppose one of these producers (indexed by $i = b$) is aware that her input demand choices influence market prices.[4] This producer is endowed with $l_b^0 = \theta L$ units of land, where $\theta \in [0, 1]$ is a measure of the degree of initial concentration in land ownership. For simplicity I refer to farmer b as the *landlord* throughout the analysis, despite the fact that having market power is independent of the size of the firm, and that the model allows for this agent to have low or even no initial land endowments. Analogously, all agents that do not exert market power and maximize their profits taking prices as given are considered to be part of the *competitive fringe*.

Occupational Choice: Individuals choose whether to work in agriculture as farmers or as wage workers outside agriculture. Individuals working outside agriculture cannot own any land, and earn income:

$$\pi_i^{out} = w^{out} s_i^{out} n_i^0 + r l_i^0,$$

where the wage in the non-agricultural sector, w^{out}, is set exogenously.

All individuals working in agriculture as farmers have access to the same technology:

$$y_i(l_i, n_i; s_i) = s_i^{1-\gamma} \left(l_i^{\alpha} n_i^{1-\alpha} \right)^{\gamma},$$

where $\gamma \in (0, 1)$ is the span-of-control parameter, $\alpha \in (0, 1)$ determines the profit share of land, and l_i and n_i are respectively the quantities of land and labor employed. Profits in the agricultural

[4] This producer can also be thought of as the share of firms in the economy who have market power and that act as a perfectly collusive cartel.

sector are then:

$$\pi_i^{ag} = y_i(l_i, n_i; s_i) - r(l_i - l_i^0) - w(n_i - n_i^0),$$

and an individual then chooses to work in agriculture operating a farm if $\pi_i^{ag} \geq \pi_i^{out}$. The inclusion of this occupational choice where agents have the possibility of working outside agriculture makes the labor supply curve faced by the landlord to be upward-sloping.

Input Prices: Denote the share of individuals who choose to work in agriculture as $\mathcal{A} \equiv \{i : \pi_i^{ag} \geq \pi_i^{out}\}$. Given a land price r and a wage rate w, these price-taking producers define their optimal demand for land and labor (l_i^*, n_i^*) simultaneously. Regardless of market structure, these producers choose input demands such that each firm's marginal productivity matches input prices: $y'_{l_i} = r$ and $y'_{n_i} = w$.

For their part, market clearing conditions require that:

$$\sum_{\substack{i \in \mathcal{A} \\ i \neq b}} l_i = L - l_b \, ; \quad \sum_{\substack{i \in \mathcal{A} \\ i \neq b}} n_i = N - n_b, \tag{1.1}$$

where $N \equiv \sum_{i \in \mathcal{A}} n_i^0$ is the supply of labor in the agricultural sector, and l_b and n_b are, respectively, the amounts of land and labor demanded by the landlord.

Combining the optimality conditions with the market-clearing equations above it is possible to express input prices in terms of the demand decisions of the landlord:

$$w = (1-\alpha)\gamma \left[\frac{(L-l_b)^{\alpha\gamma}}{(N-n_b)^{1-(1-\alpha)\gamma}} \right] \left(\sum_{\substack{i \in \mathcal{A} \\ i \neq b}} s_j \right)^{1-\gamma} ; \; r = \alpha\gamma \left[\frac{(N-n_b)^{(1-\alpha)\gamma}}{(L-l_b)^{1-\alpha\gamma}} \right] \left(\sum_{\substack{i \in \mathcal{A} \\ i \neq b}} s_j \right)^{1-\gamma} . \tag{1.2}$$

Competitive Benchmark: The landlord chooses inputs that maximize the profit function:

$$\pi_b(l_b, n_b; s_b) = s_b^{1-\gamma} \left(l_b^{\alpha} n_b^{1-\alpha} \right)^{\gamma} - r(l_b - \theta L) - w n_b, \tag{1.3}$$

where, for simplicity, I assume the landlord's endowment of labor (n_b^0) to be zero.

Under perfect competition the optimality conditions of the landlord coincide with those of all other firms (i.e. $y'_{l_b} = r, y'_{n_b} = w$). Given market clearing conditions, the optimal operational scale of each firm is determined by its relative productivity:

$$l_b^{pc} = L \times \frac{s_b}{\sum\limits_{j \in \mathcal{A}} s_j}\,; \quad l_i^{pc} = L \times \frac{s_i}{\sum\limits_{j \in \mathcal{A}} s_j}, \;\; \forall i.$$

In the perfectly competitive scenario input demands are independent of initial endowments and varying levels of initial land concentration have distributive implications but do not alter the (optimal) aggregate efficiency of the economy.

Market Power: Assume for simplicity that the landlord exerts market power only when deciding how much land (l_b) to demand but acts as a price taker regarding the amount of labor (n_b) hired.[5] This assumption implies that, analogous to firms in the competitive fringe, the landlord's marginal productivity of labor matches the market wage, $y'_{n_B} = w$. Under this assumption monopsony power in labor is thus only exerted indirectly through the effect that changes in the demand for land have on the equilibrium wage.

When deciding on its optimal demand for land the firm with market power will take into account the indirect effects on input prices. In this scenario the first-order condition for land stemming from the profit function in equation (1.3) is:

$$y'_{l_b} = r\left(1 + \left(1 - \frac{\theta L}{l_b}\right)\varepsilon_{l_b r}^{-1}\right) + \frac{w n_b}{l_b}\varepsilon_{l_b w}^{-1}, \tag{1.4}$$

where $\varepsilon_{l_b r}^{-1} \equiv \frac{\partial ln(r)}{\partial ln(l_b)}$ and $\varepsilon_{l_b w}^{-1} \equiv \frac{\partial ln(w)}{\partial ln(l_b)}$ are, respectively, the inverse elasticity of land supply with respect to land prices and wages. Following the expression for input prices in equation (1.2), these

[5] The assumption that market power is exerted in only one of the two input decision simplifies the analysis but is not necessary for the model's main results to hold. The same qualitative patterns are observed when allowing the farmer with market power to consider how her own input demands affect input prices for both land *and* labor decisions.

elasticities are also a function of the landlord's net demand for land:

$$\varepsilon_{l_b r}^{-1} = (1-\alpha\gamma)\frac{l_b}{L-l_b} \geq 0 \ , \quad \varepsilon_{l_b w}^{-1} = -\alpha\gamma\frac{l_b}{L-l_b} \leq 0.$$

Equation (1.4) implicitly defines the optimal demand for land (l_b^*) by a firm exerting market power. It is possible to show using this expression that *i)* $l_b^*|_{\theta=0} < l_b^{pc}$: under market power, if the landholder has no initial land endowment her optimal farm size will be below perfect-competition levels, and that *ii)* $\frac{\partial l_b^*}{\partial \theta} > 0$: under market power, the landlord's demand for land is monotonically increasing with respect to her initial endowment.[6]

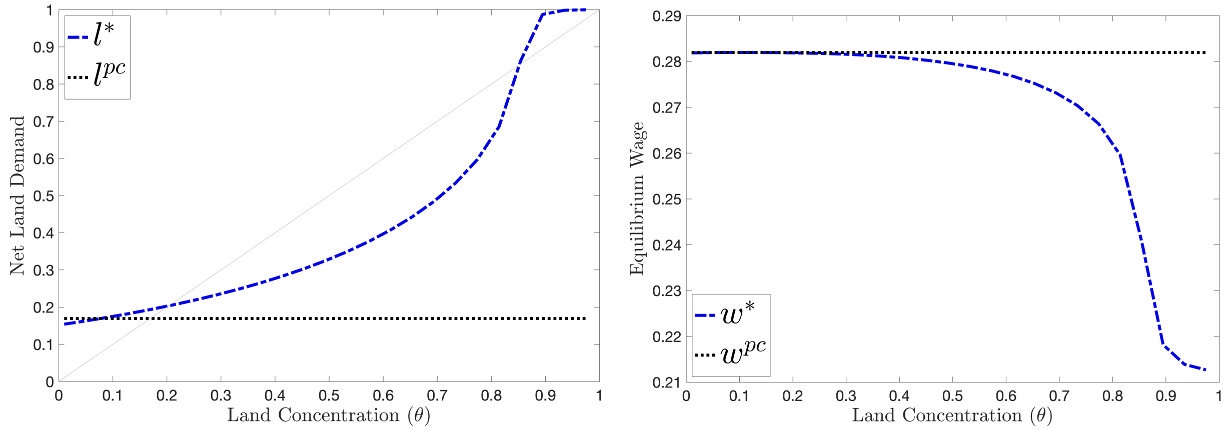
The first result mirrors a standard result in the literature studying monopsony in labor markets and states that, in order to mark-down wages, firms with market power but no land endowment will be smaller—employing less land and labor—than under perfect competition. Moreover, the higher the ability level of the landlord, the wider the gap between perfectly competitive labor demand and the one in the imperfect competition case. The second result illustrates the fact that the equilibrium farm size distribution is not independent of initial concentration levels, and implies that firms with market power will in general be inefficiently large (or small) with respect to the competitive benchmark case.

Equilibrium: While I am unable to derive an analytical expression for the landlord's demand for land, equation (1.4) makes it possible to find a numerical solution for this agent's optimal farm size given parameter values $\{\theta, \gamma, \alpha\}$, and a vector of farmer abilities $\{s_b, s_b^{out}, \overrightarrow{s_i}\ \overrightarrow{s_i^{out}}\}$. This solution in turn sequentially defines the landlord's optimal demand for labor, the input prices, the share of workers in the non-agricultural sector, and the input demand choices of the competitive fringe. The market equilibrium is then a tuple of input demands and prices $\{l_b^*, n_b^*, \overrightarrow{l_i^*}, \overrightarrow{n_i^*}, w, r\}$ such that all agents make optimal choices and the market clearing conditions in equation (1.1) hold.

Numerical solutions from a simulated economy for these equilibrium values as a function of the initial land concentration parameter are shown in Figure 1.2. The landlord's demand function for land at varying concentration levels (i.e. at varying amounts of her initial land endowment)

[6] Derivations for these results, and all expressions for prices, elasticities, and input demands are given in Appendix 1.9.

Figure 1.2: Land Demand and Wage as a Function of Initial Concentration - No Land Ceilings



(a) Net Land Demand (b) Wage

Notes: Equilibrium values for the landlord's net land demand based on the numerical result to equation (1.4) (left panel) and for equilibrium market wages (right panel) across the domain of initial concentration values θ. Blue dashed line: equilibrium values for market power case. Black dotted lines: equilibrium values for the perfectly competitive benchmark case. Details on the parameter values chosen are reported in Table 1.1.

gives rise to an S-shaped demand function that crosses the 45-degree line twice. This non-linearity arises due to the opposing incentives faced by a firm simultaneously exerting monopsony power in labor markets and monopoly power in land markets. At low levels of land concentration the monopsony incentives are stronger and (relative to the perfectly competitive case) the landlord operates a suboptimally small farm in order to mark-down wages. As concentration levels rise (i.e. as θ increases), the landlord's monopolist incentive to mark-up land prices by curtailing the amount of land supplied to markets becomes increasingly important. Consequently, the landlord's equilibrium farm size expands monotonically with land concentration levels. The landlord will continue to curtail the supply of land and operate an ever larger farm even beyond the point when it is utilizing all of her endowment. At high enough concentration levels the landlord's net demand for land becomes positive and she then turns to *buy* land from firms in the competitive fringe. This 'reverse-tenancy' scenario described by Conning (2003) can be explained by the fact that, once the operational scale of the landlord exceeds a certain threshold, the monopsonist incentive to keep wages depressed once again overcomes the monopolist incentive to profit from high land prices. At sufficiently high levels of initial concentration the landlord finds it optimal to buy out firms in the competitive fringe in order to suppress aggregate labor demand and keep wages low.

One implication of the model's results is that, in a dynamic framework, the economy's long-run equilibrium depends on the landlord's endowments. If initial land concentration exceeds a certain threshold, the economy will gravitate towards a high-inequality equilibrium where the landlord operates a large and inefficient estate, and where any firms entering the economy get bought up in order to maintain aggregate labor demand depressed. This dynamic constitutes a new potential explanation for the well-documented persistence of land inequality in Colombia despite the country's ambitious policy of smallholder land allocation (Sánchez and Villaveces, 2016; Faguet et al., 2020; Deininger, 1999). I return to this point in Section 1.6.6 where, using data on individual sales of government-allocated land plots, I find suggestive evidence consistent with this hypothesis.

1.3.2 The imposition of land ceilings

Suppose now that on a fraction $\psi \in (0,1)$ of land in the economy there is a restriction that impedes the accumulation of more than a fixed amount of land by any single farmer. Denote this land ceiling by $\bar{l} > 0$.

The imposition of this restriction can be modelled as separating available farmland into two distinct, perfectly substitutable, production inputs: restricted (l_R) and unrestricted (l_U) land. Given this assumption, the production technology available to farmers becomes:

$$y_i(l_{Ui}, l_{Ri}, n_i; s_i) = s_i^{1-\gamma}\left((l_{Ui} + l_{Ri})^{\alpha} n_i^{1-\alpha}\right)^{\gamma},$$

and the individual problem of each firm can be written as:

$$\max_{\{n_i, l_{Ui}, l_{Ri}\}} \pi_i^{ag} = y_i(l_{Ui}, l_{Ri}, n_i; s_i) - r_U(l_{Ui} - l_{Ui}^0) - r_R(l_{Ri} - l_{Ri}^0) - w(n_i - n_i^0), \tag{1.5}$$

subject to the constraints:

$$l_{Ri} \leq \bar{l};$$

$$l_{Ri}, l_{Ui}, n_i \geq 0,$$

where, as before, l_{Ui}^0, l_{Ri}^0, and n_i^0 are input endowments and l_{Ui}, l_{Ri}, and n_i are input demands.

As before, all individuals face an occupational choice and decide to work in agriculture only if their profits as farmers are higher than in the non-agricultural sector: $\pi_i^{ag} \geq \pi_i^{out}$. Continue denoting the set of individuals who choose to work in agriculture as $\mathcal{A} \equiv \{i : \pi_i^{ag} \geq \pi_i^{out}\}$.

For individuals choosing agriculture, note first that, since restricted and unrestricted land act as perfect substitutes, if $r_R > r_U$ the aggregate demand for unrestricted land is zero and no equilibrium exists given that land markets never clear at these prices. Market equilibrium thus requires that $r_R \leq r_U$. Then, as shown in Appendix 1.9, for any pair of input prices $\{r_R, r_U : r_R < r_U\}$, the ability level of each farmer determines which of the constraints above are binding in each individual solution to the firm's optimization problem. Depending on which constraints bind, each firm is placed in one of three distinct production regimes. There is first a set of fully unconstrained firms (denoted as firms $i \in S$) made up of farmers with low enough productivity levels that they do not find the ceiling on landholdings binding. There is a second set of mid-ability farmers who find the ceiling binding and would benefit from increasing their operational scale buying more land at price r_R but not at price r_U. The optimal choice for all farmers in this regime is thus to operate farms precisely at the ceiling ($\bar{l}$); denote this set of firms as $i \in C_1$. Finally, the set of highest-ability farmers operate the largest farms and for all the land demanded in excess of the mandated ceiling they pay the unrestricted-land price r_U. Denote this set of firms as $i \in C_2$. Figure 1.3 shows how farmers are sorted across these different regimes according to their innate ability level.

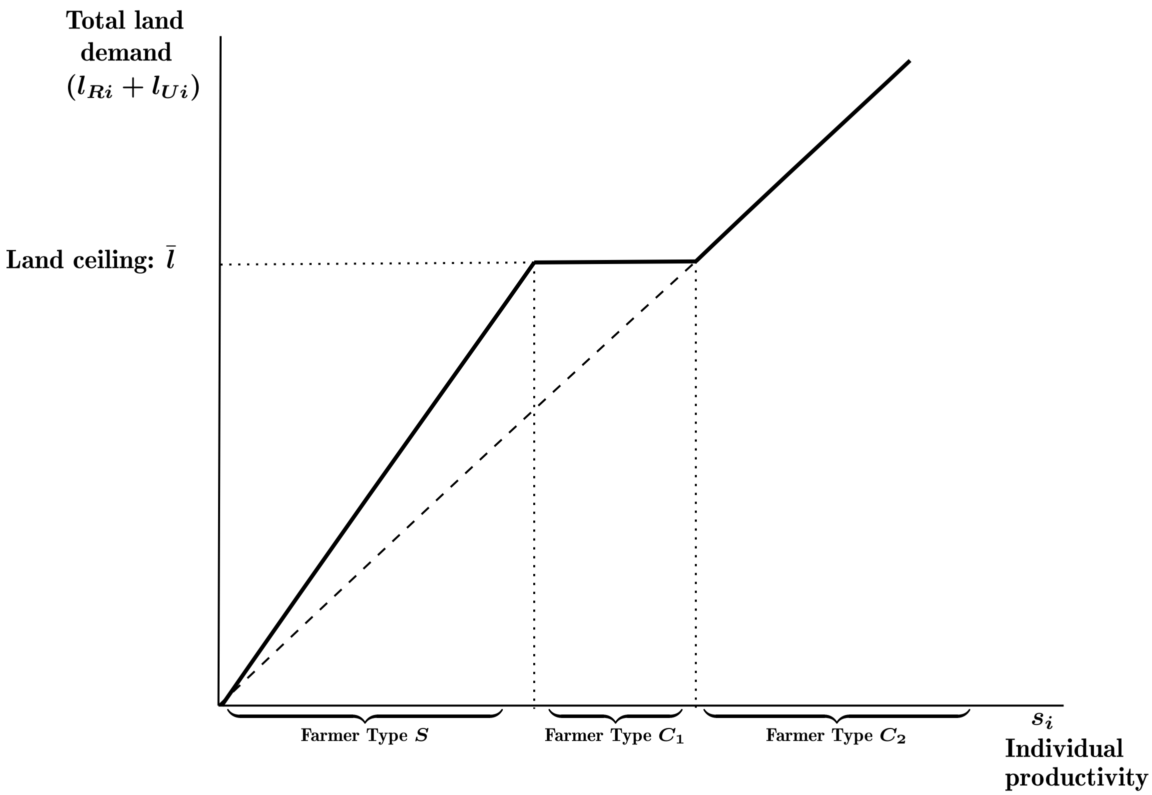
Define for convenience the auxiliary term $\rho = 1 - (1 - \alpha)\gamma$. Input demand functions in terms of prices for the three different types of firms are as follows. For farmers of type S:

$$n_i^S = s_i \left(\gamma \left(\frac{\alpha}{r_R} \right)^{\alpha\gamma} \left(\frac{1-\alpha}{w} \right)^{1-\alpha\gamma} \right)^{\frac{1}{1-\gamma}}, l_{Ri}^S = s_i \left(\gamma \left(\frac{\alpha}{r_R} \right)^{\rho} \left(\frac{1-\alpha}{w} \right)^{1-\rho} \right)^{\frac{1}{1-\gamma}}, l_{Ui}^S = 0;$$

for farmers of type C_1:

$$n_i^{C_1} = s_i^{\frac{1-\gamma}{\rho}} \left(\frac{(1-\rho)\bar{l}^{\alpha\gamma}}{w} \right)^{\frac{1}{\rho}}, \quad l_{Ri}^{C_1} = \bar{l}, \quad l_{Ui}^{C_1} = 0;$$

Figure 1.3: Total Land Demand as a Function of Individual Productivity



and for farmers of type C_2:

$$n_i^{C_2} = s_i \left(\gamma \left(\frac{\alpha}{r_U} \right)^{\alpha\gamma} \left(\frac{1-\alpha}{w} \right)^{1-\alpha\gamma} \right)^{\frac{1}{1-\gamma}},$$

$$l_{Ri}^{C_2} = \bar{l}, \quad l_{Ui}^{C_2} = s_i \left(\gamma \left(\frac{\alpha}{r_U} \right)^{\rho} \left(\frac{1-\alpha}{w} \right)^{1-\rho} \right)^{\frac{1}{1-\gamma}} - \bar{l}. \tag{1.6}$$

Recall as well that out of the total amount of land in the economy L, a fraction θ is initially owned by the firm with market power. By assumption this land is unrestricted. Out of the remaining $(1-\theta)L$ units of land, a fraction ψ is restricted and subject to the land ceiling limit. The total supply of unrestricted land is then $L_U = (1-\psi)(1-\theta)L + \theta L$ while the supply of restricted land is $L_R = \psi(1-\theta)L$. Combining these expressions with the input demand functions in equations (1.6) leads to the following set of market-clearing conditions:

for restricted land:

$$\begin{aligned} L_R = \psi(1-\theta)L &= \sum_{i \in S} l_{Ri}^S + \sum_{i \in C_1} l_{Ri}^{C_1} + \sum_{i \in C_2} l_{Ri}^{C_2} + l_{Rb} \\ &= \sum_{i \in S} s_i \left(\gamma \left(\frac{\alpha}{r_R} \right)^{\rho} \left(\frac{1-\alpha}{w} \right)^{1-\rho} \right)^{\frac{1}{1-\gamma}} + \mathcal{I}_{C_1} + \mathcal{I}_{C_2} + l_{Rb}, \end{aligned}$$

where $\mathcal{I}_{C_1} \equiv \sum\limits_{i \in C_1} \bar{l}$, and $\mathcal{I}_{C_2} \equiv \sum\limits_{i \in C_2} \bar{l}$;

for unrestricted land:

$$\begin{aligned} L_U = (1-\psi+\psi\theta)L &= \sum_{i \in S} l_{Ui}^S + \sum_{i \in C_1} l_{Ui}^{C_1} + \sum_{i \in C_2} l_{Ui}^{C_2} + l_{Ub} \\ &= \sum_{i \in C_2} s_i \left(\gamma \left(\frac{\alpha}{r_U} \right)^{\rho} \left(\frac{1-\alpha}{w} \right)^{1-\rho} \right)^{\frac{1}{1-\gamma}} - \mathcal{I}_{C_2} + l_{Ub}; \end{aligned}$$

and for labor:

$$N = \sum_{i \in S} n_i^S + \sum_{i \in C_1} n_i^{C_1} + \sum_{i \in C_2} n_i^{C_2} + n_b$$

$$= \left(\gamma\alpha^{\alpha\gamma}\left(\frac{1-\alpha}{w}\right)^{1-\alpha\gamma}\right)^{\frac{1}{1-\gamma}}\left(\frac{\sum\limits_{i\in S} s_i}{r_R^{\frac{\alpha\gamma}{1-\gamma}}} + \frac{\sum\limits_{i\in C_2} s_i}{r_U^{\frac{\alpha\gamma}{1-\gamma}}}\right) + \sum_{i\in C_1} s_i^{\frac{1-\gamma}{\rho}}\left(\frac{(1-\rho)\bar{l}^{\alpha\gamma}}{w}\right)^{\frac{1}{\rho}} + n_b, \qquad (1.7)$$

where, as before, $N \equiv \sum\limits_{i\in\mathcal{A}} n_i^0$.

The set of equations (1.7) makes it possible to compute an analytic expression for each input price that depends only on the landlord's choices and the share of firms across the different production regimes. Just as in the no-ceiling scenario described in section 1.3.1, these expressions further allow for the derivation of the inverse elasticity of the landlord's land supply with respect to prices: $\varepsilon^{-1}_{l_{Ub},r_U} \equiv \frac{\partial ln(r_U)}{\partial ln(l_{Ub})}$, $\varepsilon^{-1}_{l_{Ub},r_R} \equiv \frac{\partial ln(r_R)}{\partial ln(l_{Ub})}$, and $\varepsilon^{-1}_{l_{Ub},w} \equiv \frac{\partial ln(w)}{\partial ln(l_{Ub})}$. Maintaining the assumption that market power is only directly exerted in the decision for land (i.e., $y'_{n_b} = w$ always holds), these prices and inverse elasticities make it possible to derive an (implicit) expression for the landlord's demand for unrestricted land. In particular, this expression is obtained from the landlord's first-order condition for unrestricted land when solving the profit maximization problem described in equation (1.5):

$$y'_{l_{Ub}} = r_U\left(1 + \left(1 - \frac{\theta L}{l_{Ub}}\right)\varepsilon^{-1}_{l_{Ub},r_U}\right) + \frac{wn_b}{l_{Ub}}\varepsilon^{-1}_{l_{Ub},w} + \frac{\bar{l}r_R}{l_{Ub}}\varepsilon^{-1}_{l_{Ub},r_R}. \qquad (1.8)$$

Equilibrium: I assume throughout, and parameterize the numerical simulations such that the productivity level of the landlord is high enough to ensure that she is always of type C_2, meaning that both the restriction $l_{Rb} \leq \bar{l}$ is binding and that her demand for unrestricted land is strictly positive. With this assumption in place equation (1.8) implicitly defines the optimal land demand of the landlord, from which it is possible to find a numerical solution given an initial concentration value. As before, this demand choice then sets the landlord's optimal demand for labor, pins down input prices, and defines the input demands from price-taking firms. The market equilibrium is a tuple of input demands and prices $\{l^*_{Ub}, l^*_{Rb}, n^*_b, \overrightarrow{l^*_{Ui}}, \overrightarrow{l^*_{Ri}}, \overrightarrow{n^*_i}, w, r_U, r_R\}$ such that all agents are solving their maximization problem and the market clearing conditions in equation (1.7) hold.

Table 1.1: Model Parameterization

Parameter	Value	
Technology		
α	land-share of profits	0.7
γ	Span-of-control parameter	0.8
Skill distribution $((s, s^{out}) \sim ln\mathcal{N}(\mu, \Sigma))$		
μ	means of $(log(s), log(s^{out}))$	[0,0]
σ_s	variance of of $log(s)$	1
$\sigma_{s^{out}}$	variance of of $log(s^{out})$	1
$\sigma_{s,s^{out}}$	cov$(log(s), log(s^{out}))$	0
s_b	landlord's agricultural skill	$0.2 \times \sum_i s_i$
s_b^{out}	landlord's non-agricultural skill	$0.1 \times \frac{1}{N} \sum_i s_i^{out}$
w_i^{out}	non-agricultural wage	$s_i^{out}/15 \times N$
Land endowment $(l_{Ri}^0, l_{Ui}^0 \sim ln\mathcal{N}(\mu_l, \sigma_l))$		
σ_l	log-normal variance	1
μ_l	log-normal mean	0
Land restrictions		
θ	Share of land as landlord's endowment	$\in [0, 1]$
ψ	Share of land restricted	0.6
$\bar{l}$	Land ceiling	$1.02 \times \frac{1}{N} \sum_i l_{Ri}^0$

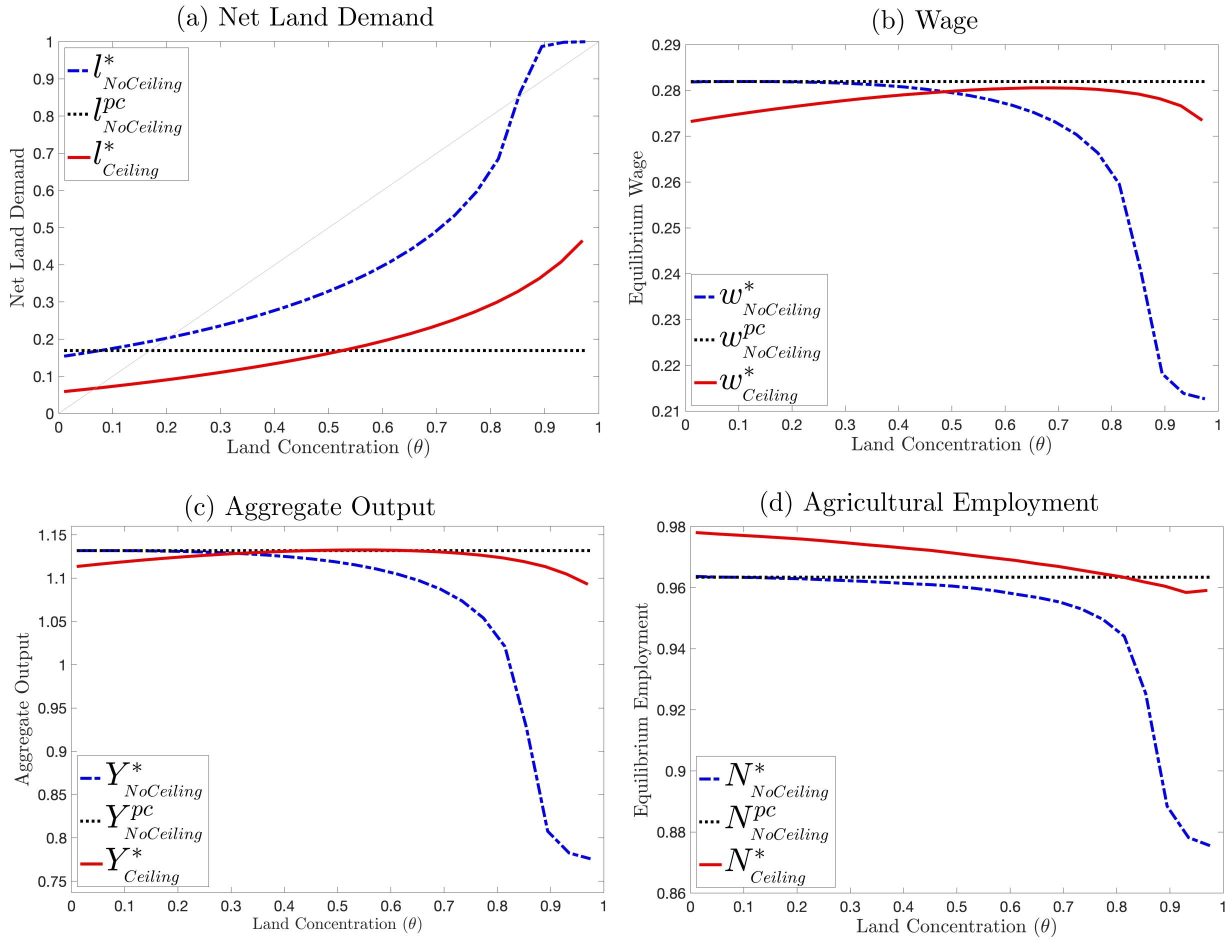
1.3.3 Numerical simulation

Computation: Table 1.1 shows the parameterization used to simulate the model. Note that all variables are defined in per capita terms so that the model is scale-invariant with respect to population size. One important issue regarding the numerical computation of the equilibrium values relates to the fact that, while all input prices and input demands (with the exception of the landlord's land demand) have analytical expressions, these expressions depend on the share of individuals choosing to work outside agriculture, as well as on the share of firms across the different production regimes S, C_1, C_2. This introduces a circular element in the definition of these objects given that the shares are ultimately defined by input prices themselves. In order to address this issue I compute the equilibrium values starting from an initial guess for the share of workers in each regime, obtain prices based on this guess, and iteratively update the shares until the difference in values across iterations approaches zero.[7]

The central goal of this numerical analysis is to illustrate how, at different levels of initial land

[7] In this sense, one additional implicit assumption of the model is that, when exerting market power, the landlord takes into account how her demand choices will affect input prices but not how those resulting prices will further affect the composition of workers across different types.

Figure 1.4: Equilibrium Outcomes as a Function of Initial Concentration - Ceiling vs. No Ceiling



Notes: Equilibrium values across the domain of initial concentration values θ. Upper left panel: landlord's net land demand based on the numerical result to equation (1.8). Upper right panel: equilibrium wages. Lower left panel: Aggregate output across all types of firms (C_1, C_2, S). Lower right panel: equilibrium employment rate in agriculture. Blue dashed lines: equilibrium values for the case of market power without land ceilings. Red solid lines: equilibrium values for case of market power with land ceilings. Black dotted lines: equilibrium values for the perfectly competitive benchmark case without land ceilings. Gray lines: equilibrium values for the perfectly competitive benchmark case with land ceilings. Details on the parameter values chosen are reported in Table 1.1.

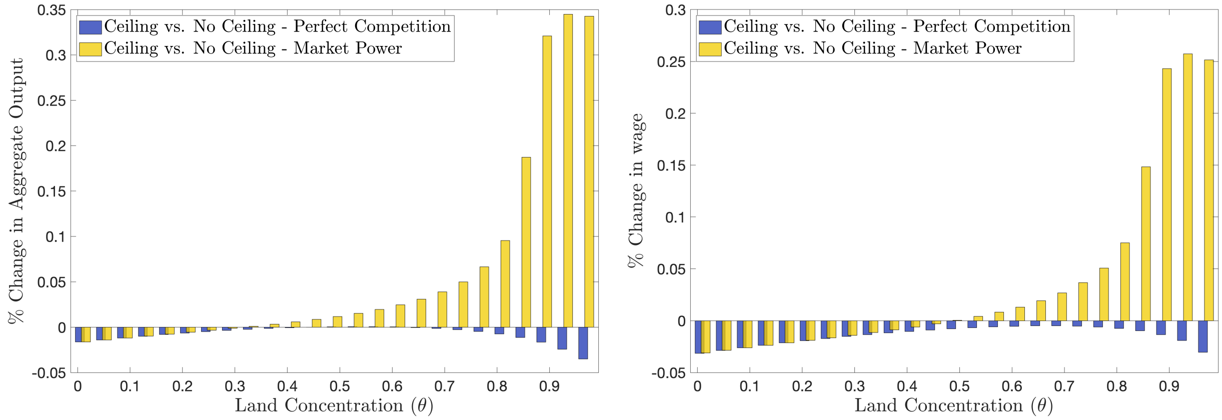
concentration, the economy's outcomes vary across the possible combinations of market structure and land market regulations: *i)* perfect competition with no land ceiling, *ii)* market-power with no land ceiling, and *iii)* perfect competition with a land ceiling.

Simulation Results: Equilibrium values for a simulated economy as a function of the initial land concentration parameter are shown in Figure 1.4. Intuitively, the introduction of a land ceiling has two countervailing effects on an economy's aggregate productivity level. On the one hand it introduces a direct distortion to the allocative efficiency of the economy by disallowing land trades that would make more productive farmers operate larger farms. In the context of the model, farmers in the production regime C_1 are the most notorious example of the misallocation produced by ceilings: farmers in this regime have varying productivity levels but the restriction on land markets leads all of them to operate equally-sized farms. Were land-market restrictions lifted, land sizes would then adjust and the marginal productivity of land would equate across all farmers in the competitive fringe. If, however, market-power distortions are also present in the economy, land-market restrictions preventing agents with market power from becoming excessively large can, potentially, increase aggregate productivity by limiting market-power effects.

As the upper-left panel of Figure 1.4 shows, the imposition of a ceiling on the amount of (restricted) land any farmer can own reduces the landlord's equilibrium farm size across the whole domain of initial concentration levels. When both land ceiling restrictions and market-power effects are at play the landlord's farm size still increases monotonically with her initial endowment, but the ceiling constrains her capacity to effectively reduce the amount of land available in the market. Given that she cannot purchase more than the maximum amount of restricted land determined by the ceiling, there is always a fraction of price-taking farms operating in the economy that the landlord is incapable of buying off. The inability to fully fend off other farms from operating—and by doing so to maintain the aggregate labor demand and market wages depressed—leads the landlord to operate a significantly smaller farm than in the no-ceiling scenario.

The reduction in the landlord's farm size caused by the imposition of ceilings has knock-on effects on input prices and aggregate productivity. Relative to the perfectly competitive case, the effect of land market ceilings on productivity and wages is unambiguously negative across the whole domain of concentration values. When market-power effects are present, however, both the magnitude and the sign of the change in productivity and wages between the unrestricted and the restricted land market cases depend on the level of initial land concentration in the economy. At low concentration levels where market-power effects are less severe, the misallocation introduced by land ceilings has a net-negative effect on productivity and wages. By contrast, at higher concentration levels the reduction in market-power distortions brought about by ceilings more than offsets its own distortionary effects and productivity is in net terms *higher* than in the unrestricted land market

Figure 1.5: Change in Outcomes across Market Structure and Restrictions



(a) Aggregate Output (b) Agricultural Wage

Notes: Log-difference in equilibrium outcome values before and after the imposition of land ceilings for the market-power scenario (yellow bars) and the perfect competition scenario(blue bars). The figure computes the difference between the solid red and dashed blue lines in the bottom left and bottom right panel of Figure 1.4 respectively. Left panel: Log-difference in aggregate output. Right panel: Log-difference in agricultural wages.

case. As illustrated in Figure 1.5, the negative productivity consequences of the imposition of land ceiling should be more severe in economies with lower initial inequality levels. I empirically test this prediction in Section 1.5.2 and find that the results are consistent with the model's predictions.

Imposing land market restrictions can have simultaneous and opposing effects on productivity and wages when concentration levels are neither too high nor too low. Figure 1.5 illustrates how the effect of land ceilings on both productivity and wages is unambiguously negative (even under imperfect competition) when initial land concentration levels are low. Conversely, at sufficiently high inequality levels the imposition of land ceilings can have a positive effect on both outcomes when market power is being exerted. However—as illustrated in Figure 1.9 in Appendix 1.8—at mid-concentration levels, land ceilings might cause increases in one outcome while leading to decreases in the other. This feature of the model is driven by the upward-sloping labor supply curve faced by the landlord, which is in turn produced by the existence of the non-agricultural sector outside option.[8] The theoretical possibility of land ceilings having opposing effects in productivity and in wages is in line with the empirical results shown in Section 1.6. I turn now to describe the data used for these estimations.

[8] A simpler model with no occupational choice—where the labor supply curve is infinitely inelastic—will also yields as a result ambiguous effects of land-market restrictions on productivity and wages that depend on initial concentration levels. This single-sector model will not, however, allow for the effect on wages and productivity to go in opposite directions at any given level of land concentration.

1.4 Data

The empirical goal of this paper is to estimate the effect that varying levels of land-market restrictions have on the agricultural sector of local economies. As I detail in Section 1.5, the main explanatory variable I use to address this question exploits the variation produced by the combination of differences in ceiling heights and the share of land in each municipality that was effectively subject to the restriction. To measure aggregate agricultural productivity I use a newly-built dataset containing information on area planted and quantity produced at the crop×municipality×year level. I aggregate and geographically link this data to 6 additional micro-level datasets to build a yearly municipality level panel. Assembling these different data sources allows me to estimate how variation in market restriction levels affected *i)* the number and type of land transactions held, *ii)* inequality in land ownership, *iii)* aggregate agricultural productivity, and *iv)* labor market conditions in agriculture. This section describes each of these datasets.

Share of municipal land restricted: To measure the fraction of land in each municipality subject to the ceiling restriction, I use the dataset from the System of Information for Rural Development (SIDER), currently mantained by the National Land Agency (ANT), which contains information on the date, area, location (at the municipal level), and recipient of every public land (*baldío*) allocation made by the Colombian government since 1900 until 2013. The dataset consists of 503,000 allocations made across 1,031 municipalities, adding up to 19 million hectares of land; about 49% of all farmland held by private individuals in 2010. A growing number of studies (e.g. Albertus and Kaplan (2013); Albertus (2019); Faguet et al. (2020); López-Uribe (2022)) use this source of information to study the determinants of land allocation patterns across time. Figure 1.10 in Appendix 1.8 shows the temporal variation in number of allocations and the amount of land granted by the government throughout the 1940–2012 period.

Land sales: To measure if the imposition of land ceilings did in fact affect the prevalence and type of land sales in Colombian municipalities, I use data from the National Superintendency of Notaries (SNR), the national-level agency in charge of supervising and keeping record of property transfers. This dataset allows me to observe all formal sales for the universe of land plots initially registered as a governmental allocation, i.e. those plots subject to the ceiling restriction. The data allows me to distinguish between full land transfers (when the totality of a plot is transferred to another individual), and partial transfers (when the original owner keeps a fraction of the plot). The data also records when two plots are merged and consolidated into a larger farm, in addition to dozens of other legal figures such as mortgages, evictions, or inheritances. The information in this dataset amounts to roughly 1.5 million observable transactions for the period 1960–2012, which I aggregate to a yearly municipal-level panel. To the best of my knowledge this is the first paper to use SNR data to study whether the imposition of market restriction did in fact have an effect on Colombian land-market dynamics, while Arteaga et al. (2023) also use the SNR dataset to measure the effect of weather shocks on rural land sales and the farm size distribution during the 2000–2011 period.

While very rich, the SNR dataset is subject to some important caveats. First, the information available is only for those land plots originally allocated by the government, and thus it is a selected sample of rural properties (albeit a selection of specifically those plots subject to the restriction). Knowing if the restriction had an impact on the transaction frequency of unrestricted land plots is therefore not possible. Second, while being the recipient of a land allocation implied receiving a formal property title, the registration of this title in the SNR was not automatic. Plots from owners who did not pursue this process, and therefore never finalized the formalization process, are not observed. Third, any informal sales—which, by definition, were not registered in any state agency—are not observable in the data. Gauging the magnitude of the last two omissions is difficult but, as shown in Figure 1.11 in Appendix 1.8, the aggregate number of yearly plot registrations observed in the SNR dataset follows a broadly similar pattern to the number of allocations observed in the—completely independent—SIDER dataset, suggesting that the majority of government-allocated plots were indeed formally registered.

Area of land sold - SIDER and SNR merge: The SNR transaction dataset lacks information on land area. To gauge how land purchases correlate with the original area of the plots being sold, I merge the SNR transaction-level dataset with the allocation-level SIDER dataset. In the absence of plot-level identifiers, merging these two datasets is based on characteristics such as the date and municipality where the allocation took place, and on the name (subject to spelling variations) under which each plot was registered. This challenge is close to the one faced by economic historians who seek to link the same individual in two different waves of a population census before the use of personal identification numbers became widespread. I therefore follow the approach proposed by Abramitzky et al. (2021), and tailor their ABE-JW algorithm to this context. At conservative parameter values (such that keeping a low probability of false-positive matches is prioritized), the algorithm produces a match rate between both datasets of 43%, yielding a sample of 213,001 allocations.

Land inequality: To assess changes in land inequality I use gini indices in agricultural land ownership and average farm sizes at the municipality level using information based on the national cadastre system for years 1985, 1993, and the period 2000–2005. This data consists of municipal-level aggregates of the national land registry maintained by the National Institute for Geographic Information (IGAC).[9] This registry is intended to be a census of rural property and aims to collect information on the location, size, and valuation of all plots in the country.[10] The dataset aggregates, at the municipal level, information from all privately-owned agricultural land plots across every municipality including both formally and informally owned plots, as well as government and non-government allocated ones. The cadastre system is meant to be continuously updating, and each municipality's rural-property registry should in theory be updated every five years at a minimum. In practice the frequency of cadastral updates varies significantly across municipalities, but Martínez (2023) shows that the updates are not driven by changes in municipality characteristics or local economic conditions like property booms. I exclude from the analysis information from municipalities where the number of registered properties in any given year is below the 99^{th} percentile of the distribution. The final municipal-level panel contains information on nearly 40 million hectares of privately-owned farmland across 982 municipalities.

Agricultural productivity: In order to measure agricultural productivity at the municipal level before and after the enactment of law 160, I collected and digitized hundreds of volumes of the

[9] Aggregate municipal data for the period 2000–2005 was made available by (IGAC) for the construction of the *Atlas of Rural Property*, (Ibánez et al., 2012). I thank Fabio Sanchez at Universidad de los Andes for sharing his data for years 1993 and 1985.

[10] With the exception of the *departamento* of Antioquia, and the cities of Bogotá and Cali, who conduct independent land surveys.

Evaluaciones Agropecuarias Municipales, a series of biannual publications made at least since 1980 at the request of the national government by each *departamento's* rural planning unit.[11] These volumes contain information on the area planted, area harvested and production data at the semester-crop-municipality level for a broad range of both perennial and non-perennial crops. The information was gathered by local authorities through a process called 'agricultural consensus', in which extension workers, producers, downstream supply chain participants and local officials were surveyed regarding each season's harvest. Between 1980 and 2000 these volumes were published independently by local state offices in at least 17 *departamentos* (with variation in publication frequency). While not as high-quality as an agricultural census, the methodology is the same as the one used in the modern version of the *Evaluaciones* which, since 2007, are carried out nationwide by the national ministry of agriculture. These publications are a rich source of information on Colombian agriculture and are potentially useful for a large number of research questions. As far as I am aware, however, these municipal productivity figures had not been digitized and harmonized until now.[12] This paper is the first to use a comprehensive dataset with municipality-level agricultural productivity measures for the 1988–2004 period. The dataset consists of more than 135,000 crop-semester-municipality observations organized in an (unbalanced) panel of 859 municipalities in 17 *departamentos* across 17 years. This comprises 69% of the country's total population and 76% of its rural population in 2005. To get a measure of municipal land productivity, I aggregate yields across crops using FAO's primary-crop producer prices, with which I compute yearly revenue-per-hectare values.[13]

Agricultural wages and employment: For municipal measures of employment in agriculture and of the share of population in rural areas I use the National Population Censuses of 1993 and 2005 carried out by the National Statistics Office (DANE). Given that the census does not collect

[11] Most often either the *Secretaría de Agricultura* or the *Unidad Regional de Planificación Agropecuaria* (URPA).

[12] The archival work to collect this data was carried out in (and I believe exhausted) the physical archives of the Ministry of Agriculture, the Colombian Agricultural Library (*Biblioteca Agropecuaria de Colombia*), the archive of the National Metereological Institute (IDEAM), and through several requests to *departamento*-level government offices of agriculture. I do not think, however, that the current dataset contains the totality of *Evaluaciones* published during the period, as many missing volumes are probably archived in regional governmental agricultural offices across the country. As such, this data collection effort is still very much a work in progress.

[13] FAO prices are available for Colombia starting only in 1990. To compute unit-value measures in previous years I set each crop's price at 1990 levels.

information on income or wages, I use data on earnings by agricultural workers from the National Household Surveys (ENH), a set of repeated cross-section household surveys available for the period 1990–2004 also carried out by DANE. The surveys are representative at the national and *departamento* level and, while not representative for individual municipalities, do contain a large set of rural municipalities randomly selected on each survey wave.

Given that the focus of this paper is on measuring the effect of the law on wages rather than overall income, for estimation I keep only wage laborers, aged 15 to 65, and employed in the agricultural sector, excluding self-employed individuals. For this subsample I use the survey's self-reported measure of monthly monetary income as a proxy for agricultural workers' earnings, as well as a self-reported measure of the average number of hours normally worked at the job. This results in a sample of 22,517 workers spanning 332 municipalities across the country.

Additional municipal characteristics: To measure if the enactment of land market restrictions ultimately affect rural-urban migration patterns, I compute the share of population living in the rural area of the municipality using the data available in the 'municipal panel' dataset maintained by CEDE at Universidad de los Andes, a large collection of municipal-level characteristics gathered from several administrative data sources.

Finally, as a measure of the intensity of high land concentration across municipalities during the pre-reform period, I use the 'latifundia intensity' measure reported in Lorente et al. (1985). This study conducted a land census in 1984 across the country and produced a measure on the extent of land concentration in each municipality. The measure is defined as the share of total farmland part of estates larger than 500 hectares.

Table 1.2 shows descriptive statistics for all of the variables used in the analysis. All monetary values are expressed in 2018 real Colombian pesos. The next section describes the empirical strategy I follow to estimate how the introduction of the UAF ceiling across the country affected these outcomes.

Table 1.2: Estimation Sample - Descriptive Statistics

	Observations	N. Years	Mean	Std. Dev.	Min	Max
Land ceiling (hectares)	1,088	1	66.4	201	5	2,269
Govt. allocated area in 1990 (%)	1,031	1	.218	.319	0	1
Total yearly land sales	64,818	18	21.7	37.8	0	853
Number of yearly full sales	64,818	18	15.6	28.8	0	825
Number of yearly fragmenting sales	64,818	18	5.13	12.4	0	255
Number of yearly consolidating sales	64,818	18	1.03	3.66	0	83
Average farm size (hectares)	37,186	8	31.2	106	.0631	2,790
Land ownership gini index	37,186	8	.635	.171	.0264	.972
Revenue per hectare (million COP)	41,510	17	12.1	15.2	.0131	243
Annual Corn Yield (tons/hectare)	27,772	17	2.65	2.43	.0533	110
Annual Coffee Yield (tons/hectare)	11,278	17	.968	.571	.0006	18.8
Annual Plantain Yield (tons/hectare)	16,410	17	6.82	22.5	.0085	1,130
Annual Rice Yield (tons/hectare)	4,748	17	7.63	4.63	.0437	25.6
Ag. worker monthly earnings (1000 COP)	109,459	15	779	1,232	8.03	67,159
Occupied in agriculture (%)	5,904	2	.475	.214	.0051	.913
Share of rural population (%)	5,904	2	.625	.226	.0136	.983
Latifundia Intensity in 1984 (%)	617	1	.125	.167	0	.988

Notes: Summary statistics for main dependent variables and outcomes. Column 1 indicates the number of municipality-pair observations. Column 2 indicates the number of years for which there is information available on the outcome variable. All monetary values are expressed in real 2018 Colombian pesos (COP).

1.5 Empirical Strategy

1.5.1 The effect of land market restrictions

I estimate the average impact that the imposition of landholding ceilings had on agricultural-sector outcomes in Colombia. My identification strategy uses cross-municipal variation in the stringency of land market restrictions due to differences in both ceiling height and the share of land restricted to estimate a difference-in-difference regression model before and after the enactment of law 160 in 1994. However, since ceiling heights were defined according to regional agroclimatic conditions, the standard two-way municipal and yearly fixed-effect model would not account for heterogeneity in restriction levels that is potentially correlated with differential trends across regions in regulation enforcement, productivity growth, and other time-varying sources of omitted variable bias.

As an example, municipalities in the country's peripheral, more sparsely populated eastern region were assigned higher ceilings due to their perceived lower land quality. If underlying agricultural productivity growth rates during the period of study in this region were lower relative to the rest of the country (for example due to the intensifying armed conflict), the cross-region comparison produced by the standard two-way fixed-effect approach would estimate a positively-biased relationship between the stringency of land restrictions and agricultural productivity.

With this in mind, my preferred estimation approach uses variation in market restriction levels only between pairs of contiguous municipalities that straddle an 'homogeneous zone' border, across which ceiling heights vary by decree. The underlying assumption for this approach is that endogenous municipal characteristics correlated with outcomes (e.g., changes in land quality, weather patterns, or transport costs) vary smoothly across municipal borders while only restriction levels jump discretely within neighboring municipality pairs. By comparing outcomes only within pairs of municipalities at opposite sides of a border, which are arguably more similar to each other than any two pair of municipalities chosen at random, this specification addresses the heterogeneity in unobserved time-varying trends that could bias standard TWFE estimates. This estimation strategy is similar to the approach of studies that evaluate the effect of state-level policies in the U.S. by comparing outcomes in county pairs located across state boundary lines (e.g. Dube et al. (2010); Cortés et al. (2022)).

Formally, let m index municipalities, t years, and p neighboring municipality-pairs. Let $y_{m,p,t}$ denote the outcome for municipality m belonging to pair p, and let C_m denote municipal land ceiling height. Given that lower ceilings represent more restricted land markets, I define the restriction level variable as the reciprocal of ceiling height: $R_m \equiv 1/C_m$. Let the share of municipal agricultural land allocated by the government before the enactment of the law be denoted by S_{m,t^0}.[14] I then run the following OLS regression:

$$y_{m,p,t} = \beta \left(R_m \times S_{m,t^0} \times T\right) + \alpha_1 \left(R_m \times T\right) + \alpha_2 \left(S_{m,t^0} \times T\right) + \phi_m + \kappa_{p,t} + \varepsilon_{m,p,t} \tag{1.9}$$

where $R_m \times S_{m,t^0}$ is the municipal degree of restriction stringency, and is the treatment variable of interest. The indicator variable T denotes post-reform time periods (i.e. $T = \mathbf{1}(t \geq 1994)$), while

[14] I compute this share as the amount of land allocated in each municipality over total farm land in the municipality according to the 2014 National Agricultural Census. I define the 'pre-reform' period to include all government allocations made up to the year 1990 (i.e., $t^0 = 1990$), but moving this cutoff one or two years either back or forward has very little impact on the estimation results.

ϕ_m represents municipality fixed effects, and $\kappa_{p,t}$ represents neighboring municipality-pair×year fixed effects. The municipality fixed effect absorbs terms such as R_m or S_{m,t^0}, along with any other characteristics that do not vary within municipality, while the contiguous municipality-pair×year fixed effects control for any unobserved time-varying shocks that occur at a level broader than the municipality pair.[15] Within a municipality pair, the inclusion of coefficients α_1 and α_2 controls for heterogeneous post-treatment trends correlated respectively to ceiling height and share of land restricted. The coefficient of interest β is the effect of the market-restriction treatment jointly defined by ceiling height and fraction of farmland restricted, and is identified under the assumption that, conditional on the heterogeneous trends and on the set of fixed effects, the treatment variable is uncorrelated with any remaining unobserved shocks in the error term. An omitted variable confounding estimates obtained from equation (1.9) would have to vary across time and within municipality pairs, differentially affecting municipalities with more stringent market frictions.

In order to test whether the identification assumption is threatened by the existence of pretrends leading up to the enactment of the law, as well as to evaluate the persistency of the estimated treatment effects across time, I estimate an event study of the form:

$$y_{p,m,t} = \sum_{\substack{h=-j \\ h\neq -1}}^{J} \beta_h (R_m \times S_{m,t^0} \times \tau_h) + \alpha_1 (R_m \times T) + \alpha_2 \left(S_{m,t^0} \times T\right) + \phi_m + \kappa_{p,t} + \epsilon_{p,m,t}, \quad (1.10)$$

where τ_h is an indicator function such that $\tau_h = \mathbf{1}\{t - 1994 = h\}$.

Note that under the contiguous municipality-pair specification each municipality can have more than one adjacent neighbor across homogeneous zones, and can therefore be part of more than one municipality pair. This implies that in a single year a municipality might appear multiple times in the estimation sample. For this reason, all regressions are weighted by the inverse number of pairs to which each municipality belongs to. Conversely, municipalities that only share boundaries with other municipalities subject to the same restriction level are not included in the sample since the

[15] A mid-point between the standard time and municipality two-way fixed-effect approach and the neighboring municipality-pair estimation is to estimate a regression with municipality and *departamento*-year fixed effects that control for unobserved time-varying trends in municipalities across different *departamentos*. Results in tables 1.11 to 1.14 in Appendix 1.8 show results for regressions with these sets of fixed effects as well as for the more standard time and municipality fixed-effect approach. In general, results across the three sets of fixed effects are similar, suggesting that the potential omitted variable bias from unobserved cross-region trends appears not to be large.

chosen estimation approach leaves no variation left to exploit from these observations. Figure 1.12 in Appendix 1.8 shows the geographical distribution of the resulting sample of municipalities on which the estimation is carried out.

Inference: Standard errors from equations (1.9) and (1.10) are potentially subject to bias due to serial correlation in municipal level outcomes, as well as to the fact that treatment is constant within an homogeneous zone.[16] Additionally, error terms are mechanically correlated across neighbor-pairs that share a common municipality given that, as discussed above, municipalities with more than one neighbor will appear repeatedly in the estimation sample.

In order to account for these potential sources of correlation, I define *departamento*-pair groupings as the set of all municipalities belonging to either one of the two *departamentos* to which a municipality-pair observation belongs to. I use two-way clustered standard errors in all regressions by *departamento* to addresses autocorrelation and common treatment across units and by *departamento*-pairs, to address correlation across neighbor-pairs that share a common municipality.

1.5.2 Heterogeneity by initial concentration levels

The theoretical framework outlined in Section 2.5 has two main predictions that can be tested in the data. First, when there is market power and relative to the unrestricted case, the net effect of imposing market restrictions on productivity depends on the economy's level of initial land concentration. At low concentration levels—where market-power distortions are minimal—the imposition of land-accumulation restrictions will simply introduce distortions that dislodge the efficient allocation of land and will unambiguously reduce aggregate output. By contrast, at high enough levels of initial concentration, while the imposition of restrictions still introduces harmful distortions, it also curtails the inefficiencies produced by imperfect competition. The model shows that in this case, if the market-power driven inefficiencies are large enough, introducing restrictions to prevent further concentration can actually increase the economy's aggregate output.

[16]The homogeneous zones that define land ceiling height are collections of municipalities that do not straddle *departamento* boundaries. Homogeneous zones are not an administrative division, and were defined *ad-hoc* for the enactment of law 160.

In order to investigate the first prediction I estimate a modified version of equation (1.9) where I introduce an additional interaction term that measures if restrictions have heterogeneous effects on agricultural productivity between municipalities with high and low initial concentration levels. Formally, let I_{m,t^0} be an indicator function for high initial land concentration in municipality m at time $t^0 < T$. I estimate:

$$
\begin{aligned}
y_{p,m,t} = {} & \beta\left(R_m \times S_{m,t^0} \times T\right) + \gamma\left(R_m \times S_{m,t^0} \times I_{m,t^0} \times T\right) \\
& + \delta_1\left(I_{m,t^0} \times S_{m,t^0} \times T\right) + \delta_2\left(I_{m,t^0} \times R_m \times T\right) \\
& + \alpha_1\left(R_m \times T\right) + \alpha_2\left(S_{m,t^0} \times T\right) + \alpha_3\left(I_{m,t^0} \times T\right) + \phi_m + \kappa_{pt} + \varepsilon_{p,m,t},
\end{aligned}
\tag{1.11}
$$

where the measure of initial land inequality comes from Lorente et al. (1985). This measure is defined as the share of total farmland in a municipality that was part of an estate larger than 500 hectares. The indicator variable I_{m,t^0} is equal to one if municipality m has a measure of latifundia above the national-level median.

1.5.3 Concentration persistence

The second testable prediction of the model is that, when no restrictions on land purchases are in place and large landholders exert market power in land and labor, public-land allocations made in high-concentration environments will tend to become concentrated faster. Since landholders in highly concentrated economies have a net-positive demand for land, purchases of plots allocated in economies with more land concentration should be more prevalent and should happen sooner.

Empirically this implies that during the pre-reform period—i.e. when there were no restrictions against land accumulation—I should observe land plots allocated in municipalities with higher land concentration levels get resold faster and more often, and that this should be accompanied by higher levels of land reconcentration across time.

To test this prediction I focus on the subsample of matched allocations across the SNR and SIDER datasets (described in Section 3.2) that took place between 1984, when the initial concentration measure was collected, and 1993, the year before the enactment of law 160. I define an 'allocation cohort' as all allocations made in a given municipality during a specific year. Within each cohort, I compute an owner-level farm size measure defined as the sum of the area of all land plots owned by the same individual in each cohort. I then track how changes in ownership across time cause cohorts to become more or less concentrated relative to the original distribution of allocations. Following Roberts and Key (2008), I measure changes in concentration with as the percent change in the area-weighted median farm.[17]

To corroborate these patterns I also run two sets of regressions that estimate the correlation between initial concentration and sale probability, and further concentration changes across time. I first run the following plot-level linear-probability regression:

$$s_{i,m,\Delta t} = \beta X_m + \delta_{d,t} + \epsilon_{i,m,t}, \tag{1.12}$$

where $s_{i,m,\Delta t}$ is a dummy variable indicating if plot i allocated in municipality m had been sold Δt years after its initial allocation: $s_{i,m,\Delta t} \equiv \mathbf{1}(Sold_{i,m,\Delta t} = 1)$, X_m is the continuous initial latifundia intensity measure from Lorente et al. (1985), and $\delta_{d,t}$ represents the inclusion of *departamento*-specific time trends.

I also run cohort-level regressions of the form:

$$\Delta A_{j,m,t} = \beta X_m + \delta_{d,t} + \eta_{j,m,t}, \tag{1.13}$$

where the outcome variable $\Delta A_{j,m,t}$ measures the change in land concentration observed within each allocation cohort j between allocation year t and time interval Δt. Standard errors in all regressions based on equations (1.12) and (1.13) are clustered at the municipality level.

[17] The area-weighted median can be thought of as the size of a farm such that half of all of the stock of land is operated by smaller farms while the other half is operated by larger farms. Results using different concentration measures such as the area-weighted mean, or the Hirschman-Herfindhal Index yield very similar results.

1.6 Results

This section examines the impact of law 160, which imposed municipality-specific limits on the amount of government-granted land any private entity is legally allowed to own. As described in Section 1.5, the stringency of the restriction on a municipality's land market will jointly depend on the height of the imposed land ceiling and on the fraction of farmland in the municipality over which such restrictions apply.

The impact of land market restrictions on land sales, land concentration, agricultural productivity, and agricultural workers' earnings is visible in the event-study graphs shown in Figure 1.6. The graphs show estimated coefficients from equation (1.10) before and after the imposition of law 160 in 1994. Supporting the identifying assumption of the difference-in-difference approach, the estimated coefficients in years before the enactment of the law are substantially smaller and almost always statistically insignificant, implying the absence of differential pre-treatment trends. The estimates show a consistent effect on all outcomes after the treatment year. The imposition of land market restrictions had persistent effects throughout (at least) the ten years following the passing of law 160, with gradual increases in the magnitude of the effect on land sales as well as on agricultural productivity. For its part, the estimated effect on workers' earnings remained consistently positive throughout the 1994–2004 period, albeit with somewhat wider confidence intervals.

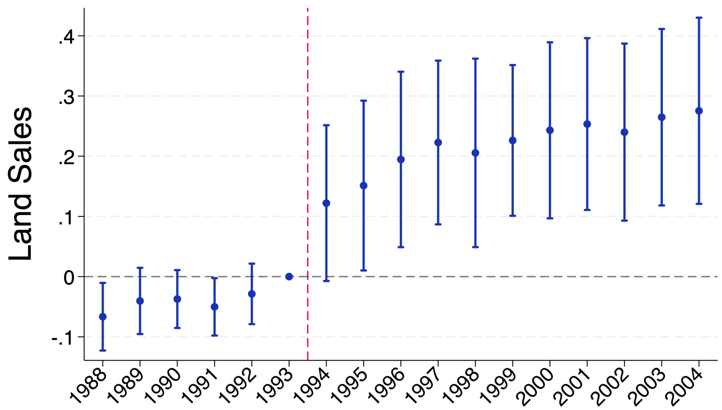
Tables tables 1.3 to 1.6 shown in the subsections below report the difference-in-difference estimates based on equation (1.9) for all outcome variables evaluated in this paper; I now turn to discuss these results.

1.6.1 Land sales

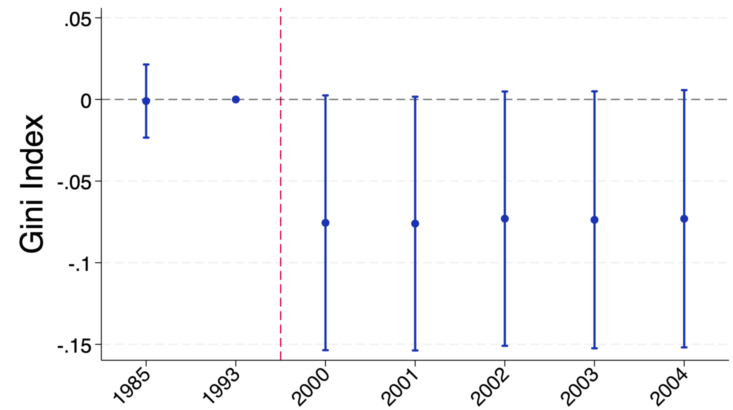
I start by looking at whether the enactment of the law had an effect on the number of land sales held in a municipality using the SNR transaction data. Looking for changes in land sales after the imposition of the law can be thought as a first-stage result evaluating whether the law was actually enforced. Recall that this data consists of government-allocated land plots, and that the number of cumulative allocations increases with time. To avoid spurious correlations due to the fact that

Figure 1.6: Event Study for Main Outcomes

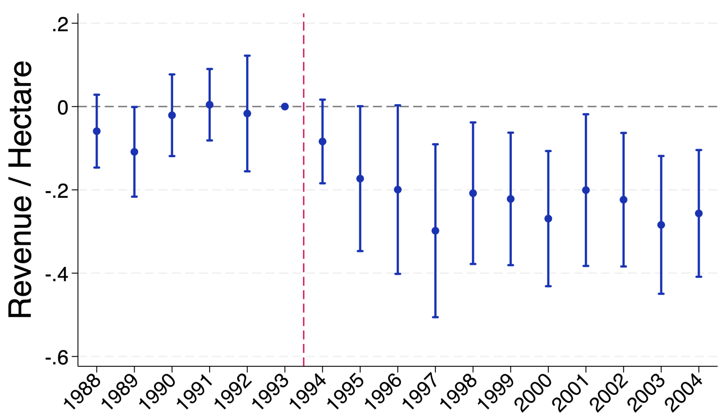
(a) Land Sales



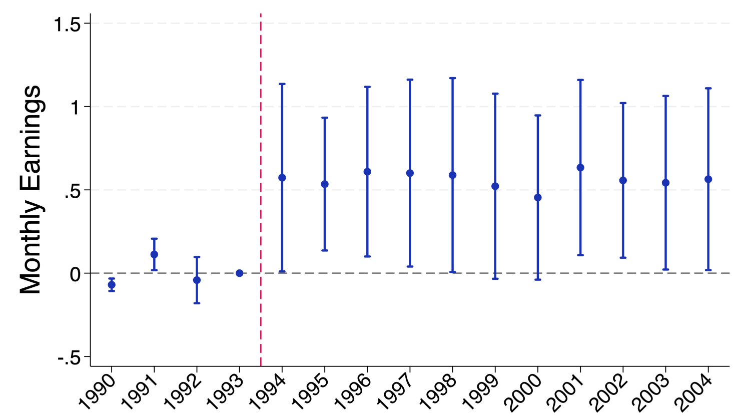
(b) Land Concentration



(c) Ag. Productivity



(d) Ag. Worker Earnings



Notes: OLS estimates of equation (1.10) for main outcome variables. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Lines around the point estimates show 95% confidence intervals for two-way clustered standard errors at the *departamento*, and at the *departamento*-pair level in parentheses.

more sales will take place as the number of plots allocated increases, these regressions include as an additional control variable the logarithm of the cumulative number of land allocations made in the municipality up to that point of time.

Table 1.3 reports results of estimating equation (1.9) on the (log) number of yearly land sales in a municipality.[18] Column 1 shows the result for the total number of sales, and columns 2-4 disaggregate by the type of sale, i.e., whether the sale transferred the property of the entire plot, whether the sale fractioned a property into smaller plots, or whether it involved merging a plot into a larger landholding.[19] Surprisingly, the imposition of limits on private landholdings *increased* the number of land sales taking place in a municipality. The point estimate indicates that a 10% increase in the composite restriction variable led on average to an increase of roughly 1.8% in the number of yearly land sales. To gauge the magnitude of this effect, consider that a municipality going from the 25^{th} to the 75^{th} percentile of restriction stringency would have 19.3% more sales every year, an increase of roughly 4.2 transactions per year at mean values.[20]

Table 1.3: Land Market Restrictions and Land Sales

		Transaction Type		
	Total Sales (1)	Full Property Transfer (2)	Fragmenting Sales (3)	Consolidating Sales (4)
$\hat{\beta}$: (log) Restriction Level × Area restricted × T	0.188***	0.191***	0.204***	-0.244**
	(0.051)	(0.046)	(0.063)	(0.092)
Observations	64,818	64,818	64,818	64,818
R^2	.956	.951	.892	.795
Mean Dep. Var.	21.708	15.612	5.133	1.025

Notes: Data from the National Superintendency of Notaries (SNR) records. Column 1 shows the effect on the aggregate number of transactions, column 2 shows the effect on full sales, column 3 shows the effect on partial sales (when only a fraction of the plot is transferred), and column 4 shows consolidation transfers. All sales variables are computed as the fraction of yearly sales in proportion to the number of cumulative government allocations at the time. All outcomes are in $\log(x+1)$ transformation. All regressions include municipality and municipality-pair-by-year fixed effects. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Two-way clustered standard errors at the at the *departamento*, and at the *departamento*-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

The positive effect in total sales was equally driven by increases in both full sales and 'partial' sales where the owner keeps a fraction of the original land plot. For their part, 'consolidation' transactions—where a plot is aggregated into a larger landholding—fell substantially:

[18] Alternatively, table 1.10 in Appendix 1.8 shows an analogous set of results where the measure of land sales is instead defined as the number of sales in proportion to the cumulative number of allocations.

[19] Formally the legal figures are *Compraventa* for full sales, *Compraventa Parcial* and *División Material* for partial sales, and *Englobe* for consolidating transfers.

[20] The difference in the log value of the treatment variable between municipalities at the 25^{th} and the 75^{th} percentiles is of |0.939 - 0.002| = 0.937 log points. The $p25/p75$ effect on percentage of sales is then $e^{0.937\times0.188} - 1 = 0.193$.

a municipality going from the 25^{th} to the 75^{th} percentile of restriction stringency would have 25.6% fewer consolidations every year after the passing of the law. I interpret these results as showing that *i)* the enactment of law 160 did have a tangible effect on land-market dynamics across Colombian municipalities, and *ii)* the simultaneous increases in fragmenting sales and decreases in consolidating sales suggest that this effect was concentrated in driving rural property towards smaller, more fragmented farms. While I cannot observe how the area of the plots transacted changed before and after the law, the results on the number and type of transactions seem to suggest a substantial number of landholders responded to the new regulations by adjusting the size of their holdings and in that way fall within the prescribed area limits.

1.6.2 Land concentration

If the imposition of the law led a large enough number of owners to adjust—at least nominally—the size of their holdings in order to remain below ceiling, this reshuffling of property could produce aggregate changes in the overall farm size distribution in a municipality. Reducing land inequality and preventing land concentration were explicit goals of law 160 but, to my knowledge, no quantitative evaluation on the effect of the law has been carried out to this date. I investigate this question using data from the National Cadastre registry from IGAC, which has pre-treatment municipal-level information on average farm sizes and municipal land gini indices for 1985 and 1993, and post-treatment indices starting in 2000. Results for these outcomes are reported in Table 1.4. Column 1 shows the enactment of law 160 had no statistically discernible impact on average farm sizes, and a negative but small impact on land concentration. The point estimate in column 2 implies that the gini index in a municipality going from the 25^{th} to the 75^{th} percentile of restriction stringency would be reduced by 7%, or 0.045 points at the mean value. This decrease amounts only to a quarter of a standard deviation in the distribution of gini indices across municipalities.

The relatively small magnitude of this effect could be in part caused by the possibility that any reductions in concentration due to the law were offset by concurrent increases in the consolidation patterns of unrestricted properties. Recall that while the law only applied to plots originally granted by the government, the national land registry includes information on all privately-owned land plots,

Table 1.4: Land Market Restrictions and Farm Size

	Average Farm Size	Land Gini
	(1)	(2)
$\hat{\beta}$: (log) Restriction Level × Area restricted × T	-0.040	-0.074*
	(0.084)	(0.039)
Observations	37,186	31,774
R^2	.99	.958
Mean Dep. Var.	31.17	.637

Notes: Data from the National Land Registry (*Catastro Nacional*) maintained by the National Geographical Office (IGAC). All outcome variables are in logarithms. All regressions include *departamento* and municipality-pair-by-year fixed effects. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Two-way clustered standard errors at the *departamento*, and at the *departamento*-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

regardless of their origin. Assessing whether individuals did in fact respond to the imposition of the law by substituting restricted for unrestricted land purchases would, however, require a more extensive dataset with information on land transactions for all types of plots. Regardless of the mechanism, these estimates show that law 160 appears to have been relatively ineffective in its goal of reducing land inequality in rural Colombia.[21]

1.6.3 Agricultural productivity

I now turn to assessing the impact of land market restrictions on agricultural land productivity. As described in Section 3.2, for this outcome I use data from the *Evaluaciones Municipales* which contain information on crop-specific yields. I aggregate yields across crops using information on farm-gate prices reported by FAO, and compute a measure of average revenue per unit of land at the municipality-year level expressed in (log) real million Colombian pesos per hectare. I also compute yields at the crop level expressed in (log) tons per hectare for the four most common crops (in terms of area planted) in this dataset during the pre-reform period 1988–1993.

Table 1.5 reports the difference-in-difference estimates from equation (1.9) for the outcomes described above. The result in column 1 shows that, in aggregate, stricter land market restriction levels led to a substantial decrease in the agricultural productivity of Colombian municipalities. On average, a municipality going from the 25^{th} to the 75^{th} percentile of restriction stringency

[21] The theoretical framework described in Section 2.5 suggests that the imposition of restrictions on a fraction of available farmland would in fact lead more skilled farmers to operate larger farms through the purchase of relatively larger amounts of unrestricted land. The model in turn predicts restrictions would create a wedge between unrestricted and restricted land prices. Unfortunately, it is not possible to empirically verify this prediction given that data on rural land prices has never been systematically collected in the country.

Table 1.5: Land Market Restrictions and Agricultural Productivity

		Yield (Tons/Hectare)			
	Revenue per Hectare	Corn	Coffee	Plantain	Rice
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}$: (log) Restriction Level × Area restricted × T	-0.235**	-0.160**	0.233***	0.202***	0.158
	(0.080)	(0.062)	(0.031)	(0.064)	(0.165)
Observations	41,510	27,772	11,278	16,410	4,748
R^2	.911	.911	.796	.857	.956
Mean Dep. Var.	12.113	2.652	.968	6.82	7.63

Notes: Data from the *Evaluaciones Agropecuarias Municipales*. Outcome in column 1 in log million Colombian pesos. Outcomes in columns 2-6 in log tons per hectare. All regressions include municipality and municipality-pair-by-year fixed effects. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Two-way clustered standard errors at the *departamento*, and at the *departamento*-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

would have a 24.6% reduction in revenue per hectare. At mean values this amounts to 2.9 million Colombian pesos at 2018 prices (≈ 980 U.S. dollars in 2018).

For individual crop yields, the imposition of restrictions led to statistically significant decreases in corn yields (the most common crop, in terms of area planted, in the country), but to increases in both coffee and plantain yields, two crops grown almost exclusively in smallholder farms with low capital-labor ratios. These effects appear consistent with the idea that market restrictions moved production towards smaller farms where crops better suited to profit from economies of scale are less productive.

I rule out the possibility that the observed effect on crop yields is caused by changes in crop composition rather than productivity per unit of land. Results reported in Table 1.15 in Appendix 1.8 show that land market restrictions had no statistically significant effect in the growth of area planted for any of the four crops examined. Moreover, the signs of the coefficients for corn and coffee are the same as the coefficients for yield estimates, implying that changes in the quantity produced of these crops must have been relatively larger—and in the same direction as yields—than any changes in area.

1.6.4 Agricultural labor

Finally, I look at the effect of land-market restrictions on the earnings of agricultural workers (both wage laborers and self-employed), on the share of agricultural employment, and on the share of population living in the rural area of a municipality. Results for these estimations are reported in Table 1.6. The estimate in column 1 shows that higher land market restriction levels led to a increase in monthly earnings for workers in the agricultural sector. On average, a worker in a municipality going from the 25^{th} to the 75^{th} percentile of restriction stringency would have a 68% increase in monthly earnings. At mean values this amounts to 546 thousand Colombian pesos at 2018 prices ($\approx$ 185 U.S. dollars in 2018). Regarding employment rates, this same increase in restriction levels would cause a 23% increase in the share of workers employed in agriculture, a rise of roughly 10 percentage points at mean values. These results are consistent with the findings of Emran and Shilpi (2020), who evaluate the impact of a prohibition on land sales in Sri Lanka and show that this restriction increased agricultural employment and wages. These authors argue that land market restrictions (who in the context they study additionally entailed the imposition of bans on rentals and mortgages) curtailed the structural transformation process in more heavily restricted areas. The estimate reported in column 3, however, shows that for the Colombian context there is no evidence of restrictions having increased rural population growth. Additionally, Table 1.16 in Appendix 1.8 shows that the increase in workers' monthly earnings is not driven by intensive-margin changes in the amount of weekly hours worked, and are not sensitive to alternative definitions of the earnings measure or to the definition of the population in the workforce.

Table 1.6: Land Market Restrictions and Labor Market Outcomes

	Ag. Worker Earnings	% Occupied in Ag.	% Pop in Rural Area
	(1)	(2)	(3)
$\hat{\beta}$: (log) Restriction Level $\times$ Area restricted $\times$ T	0.554**	0.223*	0.019
	(0.243)	(0.128)	(0.042)
Observations	102,123	5,904	5,904
R^2	.135	.93	.988
Mean Dep. Var.	802.595	.475	.625

Notes: Data from the National Population Census and the National Household Surveys carried out by the National Statistics Office (DANE). Outcome in column 1 in log thousand Colombian pesos. All regressions include municipality and municipality-pair-by-year fixed effects. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Two-way clustered standard errors at the *departamento*, and at the *departamento*-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

The simultaneous, opposing effects on productivity and wages are consistent with an economy where large landholders have the capacity to exert market power in both land and labor input markets. In particular, the observed increase in wages after the imposition of land market restrictions is consistent with the curtailment of monopsony power exerted by large agricultural employers in a rural economy where workers have limited outside options. While directly introducing distortions of their own, limits on the total amount of farmland that can be accumulated might prevent large landowners from fending-off competitors through the expansion of their holdings.

1.6.5 Treatment heterogeneity by initial land concentration

If restrictions allow for the entry of smaller agricultural firms that demand labor from the local economy, the upward pressure on wages might lead firms with market power to substantially reduce their operational scale, increase the supply of land available and, through the effect on input prices, further promote the growth of smaller competing firms. Within a framework that allows for market power effects, the impact of imposing restrictions on land markets for both productivity and wages becomes ambiguous and the direction of the effect will ultimately depend on the initial level of land concentration in the economy.

Table 1.7: Restrictions and Productivity - Heterogeneity by Initial Land Concentration

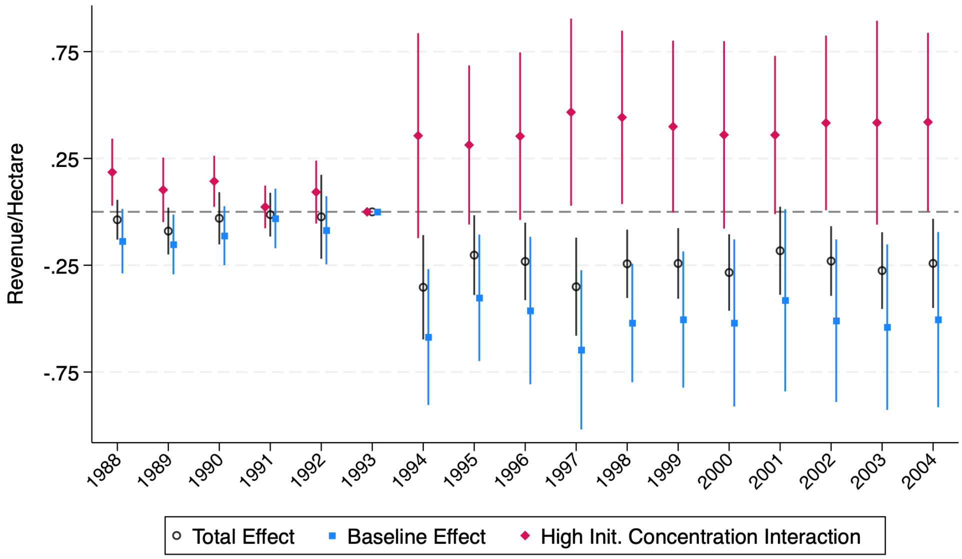
	Split Sample		Full Sample	
	Low	High		
	(1)	(2)	(3)	(4)
$\hat{\beta}$: Restriction Level × Share area restricted × T	-0.307	-0.201*	-0.217***	-0.420**
	(0.192)	(0.106)	(0.071)	(0.145)
$\hat{\gamma}$: Restriction Level × Share area restricted × T × High Init. Concentration				0.319*
				(0.170)
R^2	0.913	0.896	0.907	0.907
Observations	11,822	7,780	30,300	30,300

Notes: Productivity data from the *Evaluaciones Agropecuarias Municipales*. Initial land concentration levels from Lorente et al. (1985). Column 1 shows the estimated coefficient of running the regression specified in equation (1.9) only on the subsample of municipalities with below-median initial concentration measure. Column 2 shows the estimated coefficient of the same regression in the subsample of municipalities with above-median initial concentration. Revenue per hectare outcome in log million Colombian pesos. All regressions include municipality and municipality-pair-by-year fixed effects. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Two-way clustered standard errors at the *departamento*, and at the *departamento*-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

I test for evidence of heterogeneous effects on productivity depending on the initial level of land concentration by running the regression in equation (1.11). Results for this regression are shown in Table 1.7, which also shows the results of running the regression in equation (1.9) after splitting the sample between municipalities above or below the median latifundia intensity value.[22]

[22]The number of observations in columns 1 and 2 of Table 1.7 does not add up to the number of observations in

Figure 1.7: Agricultural Revenue per Hectare - Heterogeneity by Initial Land Concentration



Notes: OLS estimates for $\hat{\beta}$ (blue) and $\hat{\gamma}$ (red) of the event-study version of equation (1.11). Estimates in black are OLS coefficients of regression (1.9) on the full sample of municipalities with data for the initial concentration measure. Productivity data from the *Evaluaciones Agropecuarias Municipales*. Initial land concentration levels from Lorente et al. (1985). Revenue per hectare outcome in log million Colombian pesos. All regressions include municipality and municipality-pair-by-year fixed effects. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Lines around the point estimates show 95% confidence intervals for two-way clustered standard errors at the *departamento*, and at the *departamento*-pair level in parentheses.

The estimated coefficients show that, on average, the negative productivity effects of restrictions on municipalities with high concentration are less severe than in municipalities with low concentration. The interaction term in column 4 shows that while in municipalities with low concentration the effect of a 10% increase in restriction stringency leads to reduction in productivity of about 4.2%, this effect is attenuated by 3.2 percentage points in municipalities with high initial concentration levels. The heterogeneous impact of land market restrictions on productivity is also visible in the event-study graphs shown in the event-study version of equation (1.11), shown in Figure 1.7.

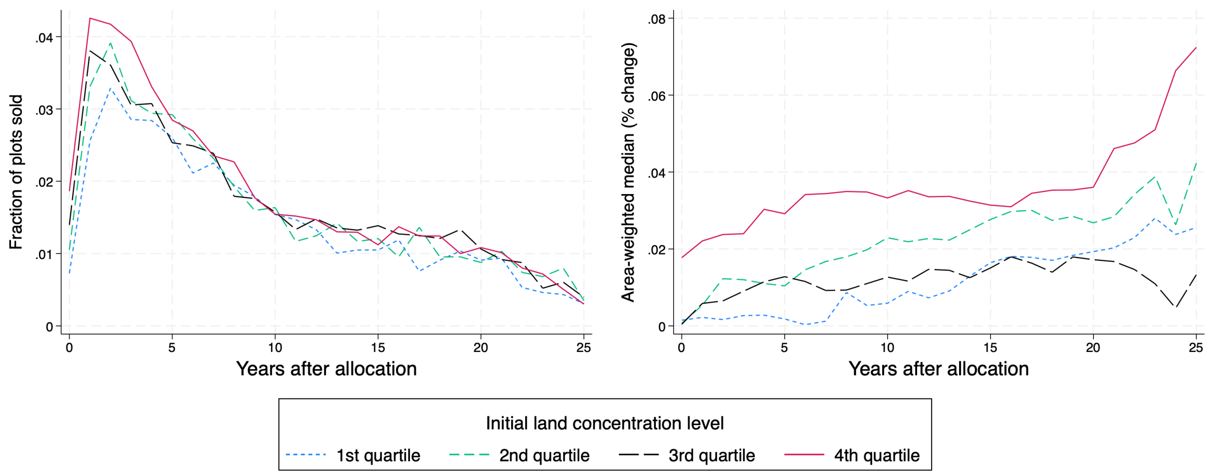
1.6.6 Concentration persistence

Figure 1.8 shows how both sales frequency and the change in concentration evolve through time after splitting the sample by quartiles of the latifundia intensity measure as described in Section 1.5.2. Consistent with the model's predictions, land plots granted in municipalities situated in the

columns 3 and 4 due to the fact that municipality-pair observations composed of one municipality with high inequality ($I_{m,t^0} = 1$), and one municipality with low inequality ($I_{m,t^0} = 0$) are excluded from estimation in both subsamples.

highest quartile of initial concentration are more likely to be sold within the first five years of allocation than land granted in municipalities in the lowest initial concentration quartile. Moreover, concentration levels increase more rapidly and remain persistently higher within allocation cohorts in municipalities with the highest prevalence of latifundia than in municipalities in the lower initial concentration quartiles.

Figure 1.8: Land Sales and Area-Weighted Median Farm Size by Initial Land Concentration



Notes: Data from the National Superintendency of Notaries (SNR) records and from from Lorente et al. (1985).

Table 1.8: Initial Concentration and Land Sales Across Time

	Land plot sold after allocation:				
	Same year	1 year later	2 years later	5 years later	10 years later
	(1)	(2)	(3)	(4)	(5)
Initial Land Concentration	0.013***	0.025*	0.024	0.030	0.041
	(0.005)	(0.014)	(0.016)	(0.024)	(0.031)
Observations	37,479	37,479	37,479	37,479	37,479
R^2	.0149	.0256	.0332	.0534	.0792

Notes: Data from the National Superintendency of Notaries (SNR) records and from from Lorente et al. (1985). Columns 1-5 show estimates of the regression described in equation (1.12) at varying intervals Δt. Estimation sample only includes allocations made between 1984 and 1993. Regressions include departamento-by-year fixed effects. Clustered standard errors at the municipality level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

The estimates reported in Tables 1.8 and 1.9—based, respectively, on equations (1.12) and (1.13)—confirm the observed correlation between initial land concentration levels in a municipality and the higher prevalence of government-allocated land to be resold and reconcentrated across time. A 1 percentage-point increase in the latifundia index measure is related to a 1.3

percentage-point increase in the probability of an allocated land plot being repurchased the same year of its allocation, with this probability increasing in time. Analogously, higher prevalence of latifundia in a municipality is correlated with the reconcentration of allocated land. Allocation cohorts in municipalities with initially higher latifundia levels tend to have faster growth in their area-weighted median farm size, with this measure of reconcentration increasing monotonically in time.

These results mirror the findings in Faguet et al. (2020), who show that land allocations did not reduce land inequality in rural economies where latifundia was initially prevalent. While the authors' explanation for this phenomenon relies in the motivation rural elites have to reconcentrate allocated land in order to retain political power, I have focused throughout this paper—following the ideas in Conning (2003)—on the possibility of an alternative explanation: imperfect competition in land and labor input markets can lead to persistently high (and inefficient) levels of land inequality.

Table 1.9: Initial Concentration and Change in Allocation Cohort's Concentration Level

	Change in area-weighted median farm size (%)				
	Same year	1 year later	2 years later	5 years later	10 years later
	(1)	(2)	(3)	(4)	(5)
Initial Land Concentration	0.056	0.067	0.063	0.076*	0.092**
	(0.041)	(0.047)	(0.047)	(0.044)	(0.047)
Observations	3,129	3,129	3,129	3,129	3,129
R^2	.0665	.0839	.0795	.0745	.0821

Notes: Data from the National Superintendency of Notaries (SNR) records and from from Lorente et al. (1985). Columns 1-5 show estimates of the regression described in equation (1.13) at varying intervals Δt. Estimation sample only includes allocations made between 1984 and 1993. Regressions include departamento-by-year fixed effects. Clustered standard errors at the municipality level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

1.7 Conclusion

This paper estimates the impact that restrictions on land markets—imposed with the aim of reducing land inequality—have on agricultural productivity, agricultural labor outcomes, and land concentration levels. I find that market restrictions permanently reduce productivity and lead only to slight reductions in land inequality. Restrictions, however, also raise the earnings of agricultural workers and the employment share in agriculture. The combination of increased workers earnings and employment is indicative of restrictions having led to a reduction in labor market power. Moreover, the heterogeneity in productivity effects across municipalities with high or low initial concentration levels are consistent with a model in which market power distortions are larger in more concentrated economies.

Conceptually, imperfect competition in land and labor markets might cause large landholders to have a willingness to pay for land that exceeds the net present value of their expected profits across time should prices be exogenous. This distorts the efficient allocation of land between producers and leads to a farm size distribution of large, unproductive estates. Agents exerting market power might find this equilibrium optimal given that withholding land from the market marks-up its price, but also because fending-off potential competitors keeps aggregate labor demand low and agricultural wages depressed. The introduction of government restrictions on land accumulation has the capacity to curtail these incentives and potentially lead the economy towards a low-concentration equilibrium. Restrictions, however, entail another source of inefficiency as they limit the reallocation of land towards productive farmers. The net effect of imposing such constraints on markets is ambiguous and depends on the relative strength of both sources of inefficiency.

The findings reported in this paper suggest that the policy of restricting land transfers has, on average, held back the efficiency of the Colombian agricultural sector, but that it has likely benefitted landless wage laborers through increases in wages. Policymakers should be aware of the potentially large distortions that market power can have on rural economies with high levels of land concentration. The decision to impose restrictions on land markets can have ambiguous effects on the economy, and there are important distributive implications related to the imposition or the elimination of such constraints in contexts where imperfect competition might be prevalent.

This paper also illustrates how market power might be a major driver of long-run land concentration. Theoretically, if land concentration levels in a rural economy are sufficiently high, unrestricted land markets might lead to an equilibrium where the existence of suboptimally large landholdings persist in time. This market-power mechanism offers a new potential explanation for Colombia's (and more generally Latin America's) well-documented persistence of land inequality and land underutilization (Deininger, 1999; Assunção, 2008a), and dispenses with traditional explanations that rely on the non-economic, cultural significance assigned to land by large landholders, or outright economic irrationality. Adjudicating between the different mechanisms that sustain high land concentration levels across time is an important question for future research.

In highly heterogeneous settings, policymakers deciding on rural land regulations are likely to face a trade-off between distinct sources of misallocation. Additional political constraints—i.e. the infeasibility of directly breaking up inefficiently large estates—might place policy makers in a situation where only second-best alternatives are possible. Future research should focus on designing and estimating the impact of innovative regulatory policies (e.g. Posner and Weyl (2017)) that are flexible in allowing markets to play a role aggregating information and allocating resources but that are as well capable of addressing potential concerns regarding the effects of imperfect competition.

Bibliography

Abramitzky, R., Boustan, L., Eriksson, K., Feigenbaum, J., and Pérez, S. (2021). Automated linking of historical data. *Journal of Economic Literature*, 59(3):865–918.

Adamopoulos, T., Brandt, L., Chen, C., Restuccia, D., and Wei, X. (2022a). Land security and mobility frictions. Technical report, National Bureau of Economic Research.

Adamopoulos, T., Brandt, L., Leight, J., and Restuccia, D. (2022b). Misallocation, selection, and productivity: A quantitative analysis with panel data from china. *Econometrica*, 90(3):1261–1282.

Adamopoulos, T. and Restuccia, D. (2020). Land reform and productivity: A quantitative analysis with micro data. *American Economic Journal: Macroeconomics*, 12(3):1–39.

Albertus, M. (2015). *Autocracy and Redistribution: The Politics of Land Reform*. Cambridge University Press, Cambridge.

Albertus, M. (2019). The effect of commodity price shocks on public lands distribution: Evidence from Colombia. *World Development*, 113:294–308.

Albertus, M. and Kaplan, O. (2013). Land Reform as a Counterinsurgency Policy: Evidence from Colombia. *Journal of Conflict Resolution*, 57(2):198–231.

Allen, D. W. (1991). Homesteading and Property Rights; Or, "How the West Was Really Won". *The Journal of Law and Economics*, 34(1):1–23.

Amodio, F. and de Roux, N. (2023). Measuring labor market power in developing countries: Evidence from colombian plants. *Jounal of Labor Economics*.

Arteaga, J., de Roux, N., Gáfaro, M., Ibáñez, A. M., and Pellegrina, H. (2023). Temperature shocks and land fragmentation: Evidence from transaction and property registry data. *Working Paper*.

Arteaga, J., Osorio, C. C., Cuéllar, D., Ibáñez, A. M., Botero, L., Murcia, M., Neva, J., and Nieto, Á. (2017). Fondo de Tierras del Acuerdo Agrario de La Habana: Estimaciones y propuestas alternativas. *Documento CEDE*, No. 41:108.

Assunção, J. (2008a). Land Reform and Land Holdings in Brazil. In Davies, J. B., editor, *Personal Wealth from a Global Perspective*. Oxford University Press, Oxford ; New York.

Assunção, J. (2008b). Rural Organization and Land Reform in Brazil: The Role of Nonagricultural Benefits of Landholding. *Economic Development and Cultural Change*, 56(4):851–870.

Banerjee, A., Gertler, P., and Ghatak, M. (2002). Empowerment and Efficiency: Tenancy Reform in West Bengal. *Journal of Political Economy*, 110(2):239–280.

Bardhan, P., Luca, M., Mookherjee, D., and Pino, F. (2014). Evolution of land distribution in west bengal 1967–2004: Role of land reform and demographic changes. *Journal of Development Economics*, 110:171–190.

Berger, D., Herkenhoff, K., and Mongey, S. (2022). Labor market power. *American Economic Review*, 112(4):1147–93.

Berry, R. A. (2017). Reflections on injustice, inequality and land conflict in Colombia. *Canadian Journal of Latin American and Caribbean Studies / Revue canadienne des études latino-américaines et caraïbes*, 42(3):277–297.

Besley, T. and Burgess, R. (2000). Land reform, poverty reduction, and growth: Evidence from india. *The Quarterly Journal of Economics*, 115(2):389–430.

Besley, T., Leight, J., Pande, R., and Rao, V. (2016). Long-run impacts of land regulation: Evidence from tenancy reform in india. *Journal of Development Economics*, 118:72–87.

Card, D. (2022). Who set your wage? *American Economic Review*, 112(4):1075–1090.

Carter, M. R. and Olinto, P. (2003). Getting Institutions “Right” for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment. *American Journal of Agricultural Economics*, 85(1):173–186.

Carter, M. R. and Zegarra, E. (2000). Land Markets and the Persistence of Rural Poverty: Post-Liberalization Policy Options. In López, R. and Valdés, A., editors, *Rural Poverty in Latin America*, pages 65–85. Palgrave Macmillan UK, London.

Chari, A., Liu, E. M., Wang, S.-Y., and Wang, Y. (2021). Property rights, land misallocation, and agricultural efficiency in china. *The Review of Economic Studies*, 88(4):1831–1862.

CNMH (2016). *Tierras y conflictos rurales: historia, políticas agrarias y protagonistas*. Centro Nacional de Memoria Histórica.

Conning, J. (2003). Latifundia Economics. *Working Paper*.

Cortés, K. R., Glover, A., and Tasci, M. (2022). The unintended consequences of employer credit check bans for labor markets. *Review of Economics and Statistics*, 104(5):997–1009.

de Janvry, A., Emerick, K., Gonzalez-Navarro, M., and Sadoulet, E. (2015). Delinking Land Rights from Land Use: Certification and Migration in Mexico. *American Economic Review*, 105(10):3125–3149.

Deininger, K. (1999). Making Negotiated Land Reform Work: Initial Experience from Colombia, Brazil and South Africa. *World Development*, 27(4):651–672.

Deininger, K. (2003). Land Markets in Developing and Transition Economies: Impact of Liberalization and Implications for Future Reform. *American Journal of Agricultural Economics*, 85(5):1217–1222.

Deininger, K., Jin, S., and Nagarajan, H. K. (2008). Efficiency and equity impacts of rural land rental restrictions: Evidence from india. *European Economic Review*, 52(5):892–918.

Dube, A., Jacobs, J., Naidu, S., and Suri, S. (2020). Monopsony in online labor markets. *American Economic Review: Insights*, 2(1):33–46.

Dube, A., Lester, T. W., and Reich, M. (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *Review of Economics and Statistics*, 92(4):945–964.

Emerick, K., Sadoulet, E., Li, Z., and Dar, M. (2022). The agricultural wage gap within rural villages. Technical report, CEPR Discussion Papers.

Emran, M. S. and Shilpi, F. (2020). Do land market restrictions hinder structural change in a rural economy? evidence from sri lanka. *The Bangladesh Development Studies*, 43(3/4):1–35.

Eslava, F. and Caicedo, F. V. (2023). Origins of latin american inequality. *Inter-American Development Bank*.

Faguet, J.-P., Sánchez, F., and Villaveces, M.-J. (2020). The perversion of public land distribution by landed elites: Power, inequality and development in Colombia. *World Development*, 136.

Felix, M. (2021). Trade, labor market concentration, and wages. *Job Market Paper*.

Galán, J. S. (2018). Tied to the Land? Intergenerational Mobility and Agrarian Reform in Colombia. *Job Market Paper*.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). Agricultural Productivity Differences across Countries. *American Economic Review*, 104(5):165–170.

Gottlieb, C. and Grobovšek, J. (2019). Communal land and agricultural productivity. *Journal of Development Economics*, 138:135–152.

Gutiérrez Sanín, F. (2019). Laberintos Institucionales: Una mirada crítica a los programas de formalización de la propiedad rural en Colombia. *Observatorio de Restitución y Regulación de Derechos de Propiedad Agraria*.

Ibánez, A., Munoz-Mora, J., and Gáfaro, M. (2012). Atlas de la distribution de la propiedad rural en colombia. *Universidad de los Andes*.

Ibáñez, A. M. and Muñoz, J. C. (2010). The Persistence of Land Concentration in Colombia: What Happened Between 2000 and 2009? In Bergsmo, M., editor, *Distrubutive justice in transitions*, number 6 in FICHL publication series. Torkel Opsahl Acad. EPubl, Oslo.

Imbert, C. and Papp, J. (2020). Costs and benefits of rural-urban migration: Evidence from india. *Journal of Development Economics*, 146:102473.

Lipsey, R. G. and Lancaster, K. (1956). The general theory of second best. *The review of economic studies*, 24(1):11–32.

López-Uribe, M. d. P. (2022). Buying off the revolution: Evidence from the colombian national peasant movement, 1957-1985. *Documento CEDE, Universidad de los Andes*, (45).

Lorente, L., Salazar, A., and Gallo, A. (1985). *Distribución de la propiedad rural en Colombia 1960-1984*. Corporación de Estudios Ganaderos y Agrícolas - CEGA.

Martinelli, P. (2014). Latifundia revisited: Market power, land inequality and agricultural efficiency. evidence from interwar italian agriculture. *Explorations in Economic History*, 54:79–106.

Martínez, L. R. (2023). Natural resource rents, local taxes, and government performance: Evidence from colombia. *The Review of Economics and Statistics*, pages 1–28.

Montero, E. (2022). Cooperative property rights and development: Evidence from land reform in el salvador. *Journal of Political Economy*, 130(1):48–93.

Naidu, S., Nyarko, Y., and Wang, S.-Y. (2016). Monopsony power in migrant labor markets: evidence from the united arab emirates. *Journal of Political Economy*, 124(6):1735–1792.

Ortiz-Becerra, K. (2021). Large-farm consolidation and structural transformation: Theory and evidence from colombia. *Job Market Paper*.

Posner, E. A. and Weyl, E. G. (2017). Property is only another name for monopoly. *Journal of Legal Analysis*, 9(1):51–123.

Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of monetary economics*, 55(2):234–250.

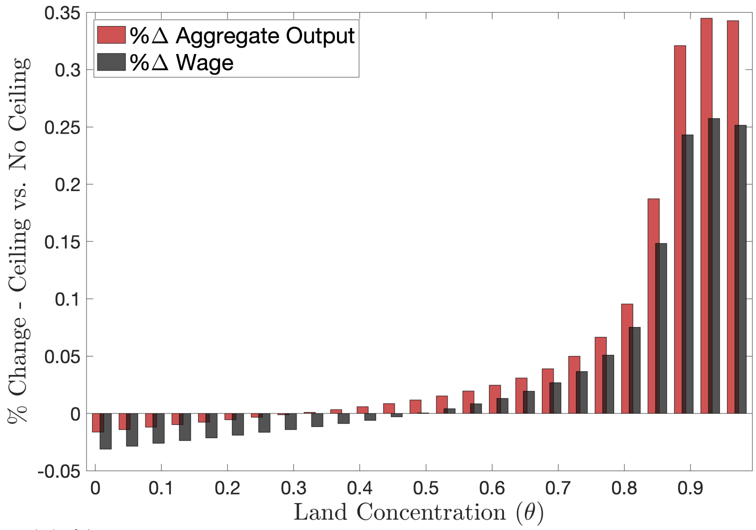
Roberts, M. J. and Key, N. (2008). Agricultural Payments and Land Concentration: A Semiparametric Spatial Regression Analysis. *American Journal of Agricultural Economics*, 90(3):627–643.

Smith, C. (2020). Land Concentration and Long-Run Development in the Frontier United States. *Job Market Paper*.

Sánchez, F. and Villaveces, M.-J. (2016). Tendencias y factores económicos y sociales asociados a la adjudicación de baldíos en Colombia, 1961-2010. In Cano, C. G., Iregui, A. M., Ramírez, M. T., and Tribín, A. M., editors, *El Desarrollo Equitativo, Competitivo y Sostenible del Sector Agropecuario en Colombia*. Banco de la República de Colombia.

1.8 Appendix A: Additional Tables and Figures

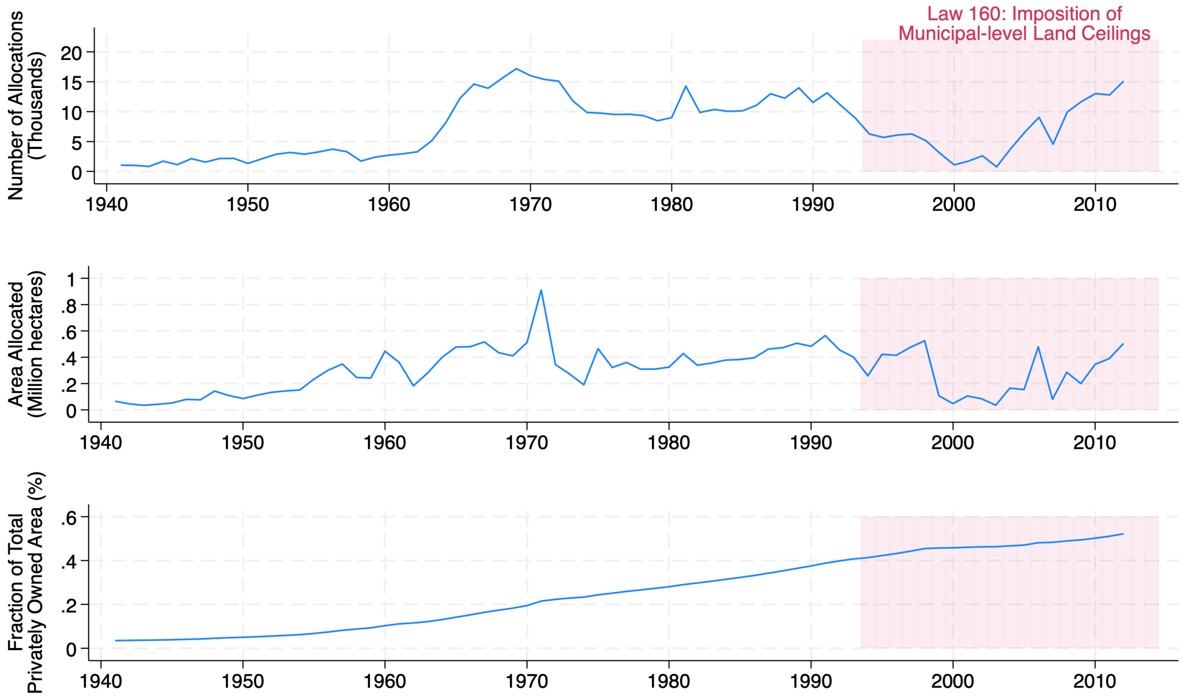
Figure 1.9: Change in Outcomes After the Imposition of Land Ceilings



(a) % Change in Output and Wages - Ceiling vs. No Ceiling

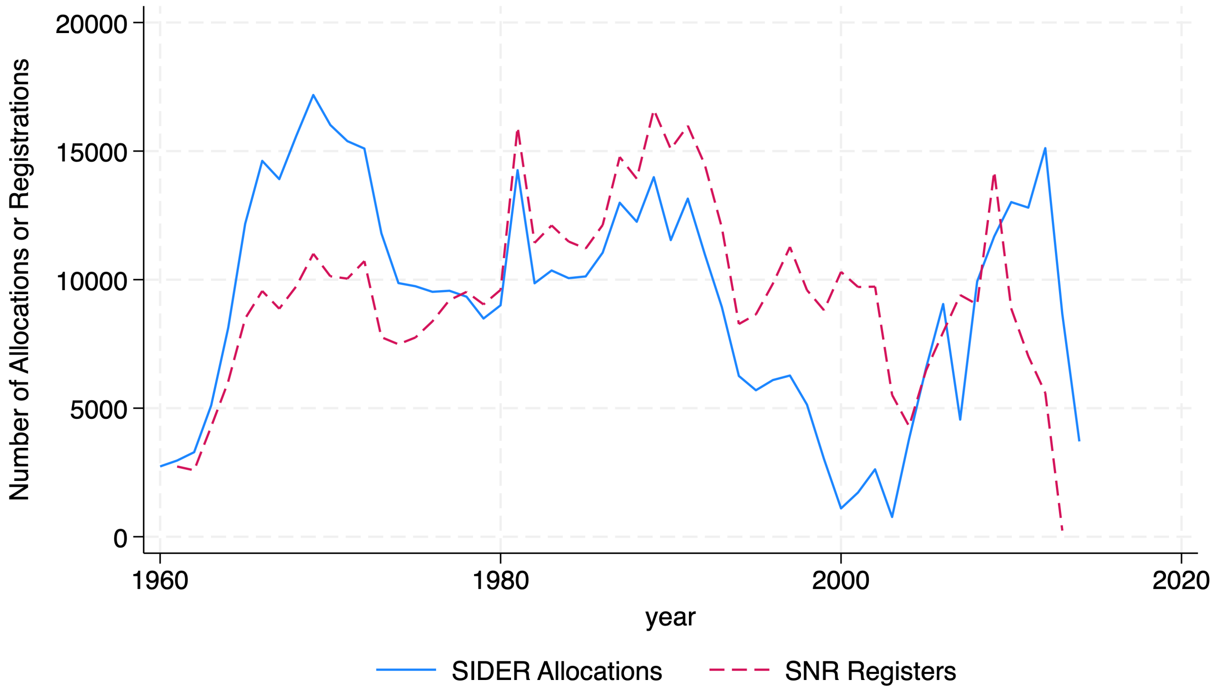
Notes: Log-difference in aggregate output and wages before and after the imposition of land ceilings when there is market power in the economy.

Figure 1.10: Number of Land Allocations and Total Area Allocated Across Time



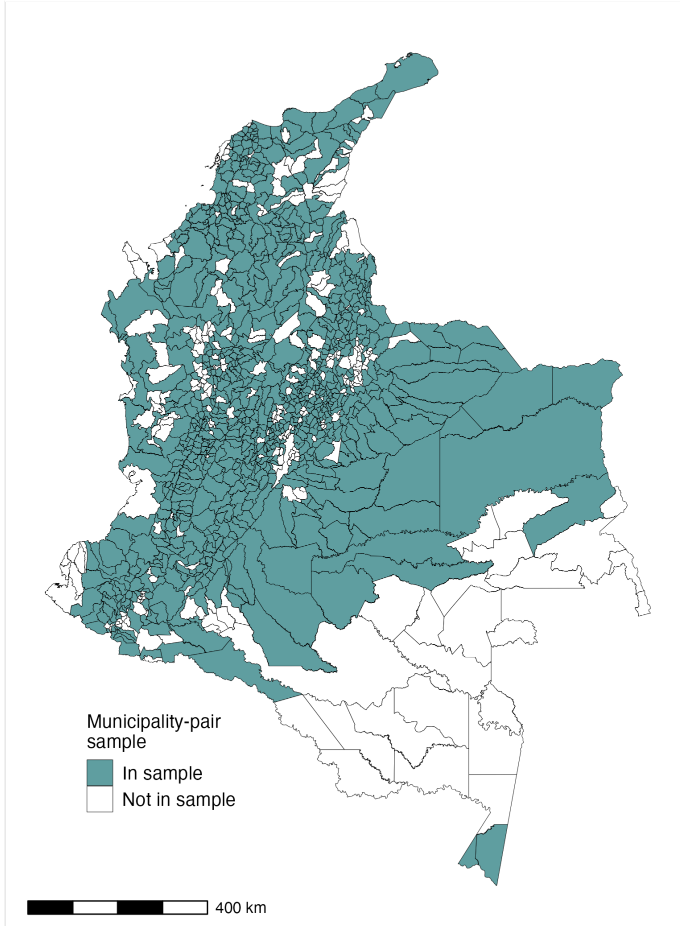
Notes: Data from the System of Information for Rural Development (SIDER)

Figure 1.11: Number of land plots allocated and number of land plots formally registered in a public notary office (1960–2010).



Notes: Data from the System of Information for Rural Development (SIDER) and from the National Superintendency of Notaries

Figure 1.12: Contiguous Municipality-Pair Sample



Notes. Ceiling height data from *INCORA* resolution 041 of 1996. Shaded municipalities have at least one neighboring municipality classified into a different 'homogeneous zone' and with a different ceiling height.

Table 1.10: Land Market Restrictions and Land Sales as Fraction of Cumulative Land Allocations

		Transaction Type		
	Total Sales	Full Property Transfer	Fragmenting Sales	Consolidating Sales
	(1)	(2)	(3)	(4)
$\hat{\beta}$: Restriction Level × Area restricted × T	0.0075**	0.0054**	0.0028	-0.0007
	(0.0032)	(0.0025)	(0.0017)	(0.0005)
Observations	64,792	64,792	64,792	64,792
R^2	.685	.684	.658	.617
Mean Dep. Var.	.041	.03	.01	.001

Notes: Data from the National Superintendency of Notaries (SNR) records. Column 1 shows the effect on the aggregate number of transactions, column 2 shows the effect on full sales, column 3 shows the effect on partial sales (when only a fraction of the plot is transferred), and column 4 shows consolidation transfers. All sales variables computed as the fraction of yearly sales in proportion to the number of cumulative government allocations at the time. All outcomes are in $\log(x+1)$ transformation. All regressions include municipality and municipality-pair-by-year fixed effects. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Two-way clustered standard errors at the *departamento*, and at the *departamento*-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 1.11: Market Restrictions and Land Sales - Varying Sets of Fixed Effects

	Number of Land Sales as Share of Allocations								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Total Sales									
$\hat{\beta}$: Restriction Level × Area restricted × T	0.164			0.137			0.188***		
	(0.109)			(0.108)			(0.051)		
$\hat{\alpha_1}$: Restriction Level × T	-0.099*	-0.048		-0.045	-0.010		-0.068**	-0.000	
	(0.049)	(0.033)		(0.074)	(0.056)		(0.030)	(0.021)	
$\hat{\alpha_2}$: Area restricted × T	0.157		-0.411***	0.142		-0.364***	0.425*		-0.272***
	(0.418)		(0.109)	(0.411)		(0.104)	(0.218)		(0.086)
R^2	0.864	0.862	0.863	0.880	0.879	0.879	0.956	0.956	0.956
Observations	24,029	24,029	24,029	23,994	23,994	23,994	64,818	64,818	64,818
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel B: Full Property Sales									
$\hat{\beta}$: Restriction Level × Area restricted × T	0.277***			0.210**			0.161***		
	(0.088)			(0.080)			(0.053)		
$\hat{\alpha_1}$: Restriction Level × T	-0.121**	-0.050		-0.038	0.011		-0.043	0.017	
	(0.048)	(0.038)		(0.071)	(0.058)		(0.033)	(0.024)	
$\hat{\alpha_2}$: Area restricted × T	0.444*		-0.250***	0.337		-0.204***	0.360		-0.241**
	(0.240)		(0.061)	(0.208)		(0.060)	(0.217)		(0.091)
R^2	0.854	0.853	0.854	0.872	0.872	0.872	0.945	0.945	0.945
Observations	19,828	19,828	19,828	19,799	19,799	19,799	83,872	83,872	83,872
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel C: Fragmenting Sales									
$\hat{\beta}$: Restriction Level × Area restricted × T	0.207			0.250***			0.169***		
	(0.126)			(0.083)			(0.061)		
$\hat{\alpha_1}$: Restriction Level × T	-0.064**	-0.003		-0.077*	-0.018		-0.016	0.053	
	(0.026)	(0.041)		(0.044)	(0.045)		(0.035)	(0.037)	
$\hat{\alpha_2}$: Area restricted × T	0.178		-0.573***	0.455		-0.469***	0.192		-0.446***
	(0.494)		(0.138)	(0.331)		(0.108)	(0.236)		(0.111)
R^2	0.719	0.716	0.719	0.755	0.754	0.755	0.889	0.888	0.889
Observations	19,828	19,828	19,828	19,799	19,799	19,799	83,872	83,872	83,872
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel D: Consolidating Sales									
$\hat{\beta}$: Restriction Level × Area restricted × T	-0.151			-0.183***			-0.250***		
	(0.098)			(0.059)			(0.090)		
$\hat{\alpha_1}$: Restriction Level × T	-0.060*	-0.113*		-0.048	-0.101**		-0.009	-0.100*	
	(0.034)	(0.059)		(0.042)	(0.038)		(0.048)	(0.049)	
$\hat{\alpha_2}$: Area restricted × T	-0.254		0.331***	-0.423**		0.262***	-0.610*		0.338***
	(0.346)		(0.104)	(0.201)		(0.077)	(0.320)		(0.081)
R^2	0.510	0.506	0.506	0.593	0.589	0.591	0.780	0.778	0.779
Observations	19,828	19,828	19,828	19,799	19,799	19,799	83,872	83,872	83,872
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X

Notes: OLS estimates of equation (1.9) for land transaction outcomes. Regressions in columns 1-6 show clustered standard errors at the *departamento* level in parentheses. Regressions in columns 7-9 show two-way clustered standard errors at the *departamento* and at the *departamento*-pair level in parentheses. Regressions in columns 7-9 are weighted by the inverse number of pairs to which each municipality belongs to. *** p<0.01, ** p<0.05, * p<0.10.

Table 1.12: Market Restrictions and Land Concentration - Varying Sets of Fixed Effects

	Average Farm Size								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Average Farm Size									
$\hat{\beta}$: Restriction Level × Area restricted × T	-0.062			-0.082			-0.040		
	(0.087)			(0.068)			(0.084)		
$\hat{\alpha_1}$: Restriction Level × T	0.102**	0.096***		0.049	0.033		0.054	0.043	
	(0.039)	(0.028)		(0.034)	(0.029)		(0.047)	(0.031)	
$\hat{\alpha_2}$: Area restricted × T	-0.323		-0.188**	-0.268		0.015	-0.151		-0.019
	(0.293)		(0.091)	(0.224)		(0.070)	(0.270)		(0.073)
R^2	0.975	0.975	0.975	0.981	0.981	0.981	0.990	0.990	0.990
Observations	12,217	12,217	12,217	12,215	12,215	12,215	37,186	37,186	37,186
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel B: Land Ownership Gini Index									
$\hat{\beta}$: Restriction Level × Area restricted × T	-0.407***			-0.235***			-0.074*		
	(0.119)			(0.070)			(0.039)		
$\hat{\alpha_1}$: Restriction Level × T	0.275***	0.215***		0.166***	0.135***		0.096***	0.078***	
	(0.063)	(0.058)		(0.036)	(0.042)		(0.027)	(0.023)	
$\hat{\alpha_2}$: Area restricted × T	-1.853***		-0.587***	-1.124***		-0.328***	-0.331*		-0.090
	(0.498)		(0.207)	(0.324)		(0.115)	(0.175)		(0.053)
Observations	12,166	12,166	12,166	12,164	12,164	12,164	31,774	31,774	31,774
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X

Notes: OLS estimates of equation (1.9) for farm size outcomes. Regressions in columns 1-6 show clustered standard errors at the *departamento* level in parentheses. Regressions in columns 7-9 show two-way clustered standard errors at the *departamento* and at the *departamento*-pair level in parentheses. Regressions in columns 7-9 are weighted by the inverse number of pairs to which each municipality belongs to. *** p<0.01, ** p<0.05, * p<0.10.

Table 1.13: Market Restrictions and Productivity - Varying Sets of Fixed Effects

	Agricultural Revenue per Hectare								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Revenue per Hectare									
$\hat{\beta}$: Restriction Level × Area restricted × T	-0.608***			-0.451***			-0.235**		
	(0.131)			(0.131)			(0.080)		
$\hat{\alpha_1}$: Restriction Level × T	0.177***	0.089		0.112***	0.055		0.021	-0.034	
	(0.043)	(0.058)		(0.037)	(0.048)		(0.046)	(0.050)	
$\hat{\alpha_2}$: Area restricted × T	-2.358***		-0.336	-1.705***		-0.147	-0.901**		-0.080
	(0.405)		(0.224)	(0.475)		(0.238)	(0.382)		(0.184)
R^2	0.729	0.726	0.725	0.769	0.768	0.768	0.911	0.911	0.911
Observations	8,281	8,281	8,281	8,281	8,281	8,281	41,510	41,510	41,510
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel B: Corn Yield per Hectare									
$\hat{\beta}$: log Restriction Level × Area restricted × T	-0.208**			-0.138			-0.160**		
	(0.089)			(0.098)			(0.062)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.006	-0.039		0.113*	0.094		0.067	0.033	
	(0.054)	(0.048)		(0.056)	(0.054)		(0.048)	(0.047)	
$\hat{\alpha_2}$: Area restricted × T	-0.800**		-0.070	-0.478		-0.032	-0.607**		-0.064
	(0.308)		(0.129)	(0.339)		(0.088)	(0.214)		(0.060)
R^2	0.714	0.713	0.713	0.775	0.775	0.773	0.911	0.911	0.911
Observations	6,396	6,396	6,396	6,396	6,396	6,396	27,772	27,772	27,772
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel C: Coffee Yield per Hectare									
$\hat{\beta}$: log Restriction Level × Area restricted × T	-0.108			-0.081			0.233***		
	(0.076)			(0.065)			(0.031)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.006	-0.022		0.084	0.048		-0.096**	-0.039	
	(0.044)	(0.037)		(0.050)	(0.041)		(0.036)	(0.034)	
$\hat{\alpha_2}$: Area restricted × T	-0.248		0.079	-0.061		0.187***	0.881***		0.136
	(0.228)		(0.094)	(0.169)		(0.055)	(0.114)		(0.096)
R^2	0.465	0.464	0.464	0.543	0.542	0.543	0.796	0.795	0.795
Observations	2,715	2,715	2,715	2,715	2,715	2,715	11,278	11,278	11,278
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel D: Plantain Yield per Hectare									
$\hat{\beta}$: log Restriction Level × Area restricted × T	-0.104			0.004			0.202***		
	(0.155)			(0.123)			(0.064)		
$\hat{\alpha_1}$: log Restriction Level × T	0.117*	0.089*		-0.032	-0.032		-0.032	0.047	
	(0.065)	(0.044)		(0.060)	(0.058)		(0.044)	(0.048)	
$\hat{\alpha_2}$: Area restricted × T	-0.394		-0.036	0.044		0.035	0.727***		0.009
	(0.599)		(0.182)	(0.477)		(0.081)	(0.215)		(0.108)
R^2	0.516	0.516	0.514	0.689	0.689	0.689	0.857	0.857	0.856
Observations	4,049	4,049	4,049	4,048	4,048	4,048	16,410	16,410	16,410
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel E: Rice Yield per Hectare									
$\hat{\beta}$: log Restriction Level × Area restricted × T	-0.036			-0.013			0.158		
	(0.171)			(0.097)			(0.165)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.081	-0.091		0.000	-0.005		-0.132	-0.048	
	(0.108)	(0.060)		(0.039)	(0.008)		(0.105)	(0.035)	
$\hat{\alpha_2}$: Area restricted × T	-0.331		-0.172	-0.106		-0.053	0.626		0.001
	(0.696)		(0.216)	(0.434)		(0.166)	(0.671)		(0.135)
R^2	0.834	0.833	0.833	0.907	0.907	0.907	0.956	0.956	0.956
Observations	1,349	1,349	1,349	1,343	1,343	1,343	4,748	4,748	4,748
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X

Notes: OLS estimates of equation (1.9) for land productivity outcomes. Regressions in columns 1-6 show clustered standard errors at the *departamento* level in parentheses. Regressions in columns 7-9 show two-way clustered standard errors at the *departamento* and at the *departamento*-pair level in parentheses. Regressions in columns 7-9 are weighted by the inverse number of pairs to which each municipality belongs to. *** p<0.01, ** p<0.05, * p<0.10.

Table 1.14: Market Restrictions and Agricultural Labor - Varying Sets of Fixed Effects

	Monthly Earnings for Agricultural Wage Workers								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Ag. Worker Earnings									
$\hat{\beta}$: log Restriction Level × Area restricted × T	0.232*			0.361**			0.554**		
	(0.120)			(0.132)			(0.243)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.056	-0.007		-0.100	-0.092*		-0.271*	-0.069	
	(0.038)	(0.042)		(0.106)	(0.049)		(0.141)	(0.084)	
$\hat{\alpha_2}$: Area restricted × T	0.810**		0.014	1.245***		0.082	1.734**		-0.046
	(0.370)		(0.123)	(0.334)		(0.144)	(0.581)		(0.237)
R^2	0.105	0.106	0.106	0.127	0.125	0.127	0.135	0.135	0.135
Observations	14,022	13,083	13,083	13,083	14,022	13,083	102,123	102,123	102,123
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel B: % Occupied in Ag.									
$\hat{\beta}$: log Restriction Level × Area restricted × T	0.016			0.190**			0.223*		
	(0.087)			(0.096)			(0.125)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.019	-0.015		-0.061	-0.028		-0.080	-0.014	
	(0.035)	(0.027)		(0.050)	(0.042)		(0.062)	(0.046)	
$\hat{\alpha_2}$: Area restricted × T	0.029		-0.003	0.589**		0.102	0.842*		-0.007
	(0.230)		(0.063)	(0.255)		(0.081)	(0.464)		(0.152)
R^2	0.849	0.849	0.849	0.861	0.860	0.861	0.930	0.929	0.929
Observations	1,446	1,446	1,446	1,444	1,444	1,444	5,904	5,904	5,904
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel C: % Pop. in Rural Area									
$\hat{\beta}$: log Restriction Level × Area restricted × T	0.037			0.067			0.019		
	(0.039)			(0.042)			(0.045)		
$\hat{\alpha_1}$: log Restriction Level × T	0.010	0.017**		-0.029**	-0.018		-0.018	-0.014	
	(0.009)	(0.007)		(0.013)	(0.012)		(0.017)	(0.014)	
$\hat{\alpha_2}$: Area restricted × T	0.134		-0.014	0.275*		0.033	0.116		0.044
	(0.136)		(0.035)	(0.154)		(0.036)	(0.162)		(0.040)
R^2	0.970	0.970	0.969	0.974	0.974	0.974	0.988	0.988	0.988
Observations	1,446	1,446	1,446	1,444	1,444	1,444	5,904	5,904	5,904
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X

Notes: OLS estimates of equation (1.9) for agricultural labor outcomes. Regressions in columns 1-6 show clustered standard errors at the *departamento* level in parentheses. Regressions in columns 7-9 show two-way clustered standard errors at the *departamento* and at the *departamento*-pair level in parentheses. Regressions in columns 7-9 are weighted by the inverse number of pairs to which each municipality belongs to. *** p<0.01, ** p<0.05, * p<0.10.

Table 1.15: Market Restrictions and Area Planted - Varying Sets of Fixed Effects

	Agricultural Revenue per Hectare								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Corn Hectares Planted									
$\hat{\beta}$: log Restriction Level × Area restricted × T	0.324**			0.277**			-0.060		
	(0.133)			(0.116)			(0.139)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.063	-0.026		-0.097	-0.072		0.054	0.024	
	(0.044)	(0.049)		(0.067)	(0.065)		(0.101)	(0.086)	
$\hat{\alpha_2}$: Area restricted × T	1.506***		0.416*	1.238***		0.298	0.172		0.370*
	(0.323)		(0.220)	(0.344)		(0.230)	(0.434)		(0.197)
R^2	0.771	0.770	0.771	0.787	0.786	0.786	0.907	0.907	0.907
Observations	6,406	6,406	6,406	6,406	6,406	6,406	27,846	27,846	27,846
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel B: Coffee Hectares Planted									
$\hat{\beta}$: log Restriction Level × Area restricted × T	-0.004			-0.053			0.087		
	(0.230)			(0.302)			(0.108)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.000	-0.006		0.049	0.047		-0.092	-0.063	
	(0.070)	(0.030)		(0.221)	(0.160)		(0.082)	(0.052)	
$\hat{\alpha_2}$: Area restricted × T	-0.069		-0.057	-0.336		-0.169	0.264		0.008
	(0.829)		(0.112)	(1.116)		(0.205)	(0.323)		(0.079)
R^2	0.922	0.922	0.922	0.942	0.942	0.942	0.979	0.979	0.979
Observations	2,729	2,729	2,729	2,729	2,729	2,729	11,320	11,320	11,320
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel C: Plantain Hectares Planted									
$\hat{\beta}$: log Restriction Level × Area restricted × T	-0.103			-0.139			-0.175*		
	(0.095)			(0.084)			(0.098)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.111	-0.139**		0.074	0.040		-0.049	-0.124**	
	(0.064)	(0.061)		(0.056)	(0.053)		(0.085)	(0.054)	
$\hat{\alpha_2}$: Area restricted × T	-0.169		0.200	-0.452		0.036	-0.539		0.107
	(0.365)		(0.152)	(0.372)		(0.105)	(0.347)		(0.144)
R^2	0.812	0.812	0.811	0.850	0.850	0.850	0.939	0.939	0.939
Observations	4,062	4,062	4,062	4,061	4,061	4,061	16,522	16,522	16,522
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X
Panel D: Rice Hectares Planted									
$\hat{\beta}$: log Restriction Level × Area restricted × T	0.104			0.068			-0.163		
	(0.178)			(0.164)			(0.229)		
$\hat{\alpha_1}$: log Restriction Level × T	-0.025	0.018		0.026	0.050		0.172	0.085	
	(0.114)	(0.057)		(0.099)	(0.044)		(0.187)	(0.077)	
$\hat{\alpha_2}$: Area restricted × T	0.219		-0.185	0.063		-0.206	-0.624		0.020
	(0.635)		(0.230)	(0.643)		(0.290)	(0.946)		(0.261)
R^2	0.853	0.853	0.853	0.880	0.880	0.880	0.944	0.944	0.944
Observations	1,355	1,355	1,355	1,349	1,349	1,349	4,768	4,768	4,768
Municipality FE	X	X	X	X	X	X	X	X	X
Year FE	X	X	X						
Departamento × Year FE				X	X	X			
Municipality-pair × Year FE							X	X	X

Notes: OLS estimates of equation (1.9) for area planted by crop. Regressions in columns 1-6 show clustered standard errors at the *departamento* level in parentheses. Regressions in columns 7-9 show two-way clustered standard errors at the *departamento* and at the *departamento*-pair level in parentheses. Regressions in columns 7-9 are weighted by the inverse number of pairs to which each municipality belongs to. *** p<0.01, ** p<0.05, * p<0.10.

Table 1.16: Effects on Agricultural Labor - Hours Worked and Alternative Measures of Earnings

	Ag. Workers' Earnings				Hours Worked	
	Monetary Income Ages 15-65	Monetary Income All ages	Total Income Ages 15-65	Total Income All ages	Hours Worked Ages 15-65	Hours Worked All ages
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}$: log Restriction Level $\times$ Area restricted $\times$ T	0.554**	0.408*	0.562**	0.495**	0.042	0.048
	(0.243)	(0.214)	(0.237)	(0.223)	(0.052)	(0.063)
Observations	102,123	109,459	104,903	112,686	103,896	111,651
R^2	.135	.133	.114	.111	.0537	.0507
Mean Dep. Var.	802.595	779.383	890.727	868.412	889.296	866.983

Notes: Data from the National Population Census and the National Household Surveys carried out by the National Statistics Office (DANE). Outcomes in columns 1-4 in log thousand Colombian pesos. Outcomes in columns 5-6 in log weekly hours. All regressions include municipality and municipality-pair-by-year fixed effects. Column 1 replicates the baseline result for agricultural workers' income in Table 1.6. All regressions weighted by the inverse number of pairs to which each municipality belongs to. Two-way clustered standard errors at the *departamento*, and at the *departamento*-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

1.9 Appendix B: Mathematical Appendix

Market power with no land ceiling: expressions for prices, elasticities, and input demands

Define for convenience the auxiliary term $\rho = 1 - (1-\alpha)\gamma$. Without land ceilings firms on the competitive fringe maximize the profit function:

$$\pi_i(l_i, n_i; s_i) = s_i^{1-\gamma}\left(l_i^{\alpha} n_i^{1-\alpha}\right)^{\gamma} - wn_i - rl_i.$$

First order conditions yield input demands in terms of prices:

$$l_i = s_i\left[\gamma\left(\frac{\alpha}{r}\right)^{\rho}\left(\frac{1-\alpha}{w}\right)^{1-\rho}\right]^{\frac{1}{1-\gamma}}, \quad n_i = s_i\left[\gamma\left(\frac{\alpha}{r}\right)^{\alpha\gamma}\left(\frac{1-\alpha}{w}\right)^{1-\alpha\gamma}\right]^{\frac{1}{1-\gamma}},$$

which, combined with market clearing conditions,

$$L = \sum_{i\neq b} l_i + l_b\,; \quad N = \sum_{i\neq b} n_i + n_b,$$

yield the expressions for input demands:

$$l_i = (L - l_b) \times \frac{s_i}{\sum\limits_{j\neq b} s_j}\,; \quad n_i = (N - n_b) \times \frac{s_i}{\sum\limits_{j\neq b} s_j}, \ \ \forall i,$$

and thus the expressions for prices and elasticities in terms of parameters and landlords' choices shown in equation (1.2):

$$w = (1-\rho)\left[\frac{(L-l_b)^{\alpha\gamma}}{(N-n_b)^{\rho}}\right]\left(\sum_{j\neq b} s_j\right)^{1-\gamma}, \quad r = \alpha\gamma\left[\frac{(N-n_b)^{(1-\rho)}}{(L-l_b)^{1-\alpha\gamma}}\right]\left(\sum_{j\neq b} s_j\right)^{1-\gamma},$$

$$\varepsilon_{l_b w}^{-1} \equiv \frac{\partial ln(w)}{\partial ln(l_b)} = -\alpha\gamma\frac{l_b}{L-l_b} \leq 0, \quad \varepsilon_{l_b r}^{-1} \equiv \frac{\partial ln(r)}{\partial ln(l_b)} = (1-\alpha\gamma)\frac{l_b}{L-l_b} \geq 0\,. \tag{1.14}$$

Regarding the landlord's optimal input demands, this agent maximizes the profit function in equation (1.3):

$$\pi_b(l_b, n_b; s_b) = s_b^{1-\gamma} \left(l_b^{\alpha} n_b^{1-\alpha}\right)^{\gamma} - r(l_b - \theta L) - w n_b.$$

The assumption that market power is only exerted when deciding how much land to demand implies that $y'_{l_B} = w$, i.e.,

$$s_b^{1-\gamma} l_b^{\alpha\gamma} n_b^{-\rho} = \left[\frac{(L - l_b)^{\alpha\gamma}}{(N - n_b)^{\rho}}\right] \left(\sum_{j \neq b} s_j\right)^{1-\gamma},$$

yields b's reaction function for labor demand in terms of land demanded:

$$\frac{N - n_b}{n_b} = \left(\frac{L - l_b}{l_b}\right)^{\frac{\gamma\alpha}{\rho}} \left(\frac{s_b}{\sum_{j \neq b} s_j}\right)^{\frac{\gamma-1}{\rho}} \tag{1.15}$$

$$\implies n_b = N \left[1 + \left(\frac{l_b}{L - l_b}\right)^{\frac{\alpha\gamma}{\rho}} \left(\frac{s_b}{\sum\limits_{j \neq b} s_j}\right)^{\frac{1-\gamma}{\rho}}\right]^{-1}.$$

By contrast, when deciding on its optimal land demand, the landlord takes into account how her demand impacts input prices. Hence, the first order condition of the profit function with respect to land implies that:

$$
\begin{aligned}
&\alpha\gamma \underbrace{s_b^{1-\gamma} n_b^{(1-\alpha)\gamma} l_b^{\alpha\gamma-1}}_{=y/l_b} = \frac{\partial r}{\partial l_b} l_b + r - \frac{\partial r}{\partial l_b}\theta L + \frac{\partial w}{\partial l_b} n_b \\
&\implies \alpha\gamma \frac{y}{l_b} = r\left(1 + \left(1 - \frac{\theta L}{l_b}\right)\varepsilon_{l_b r}^{-1}\right) + \underbrace{w n_b}_{=(1-\alpha)\gamma y} \varepsilon_{l_b w}^{-1} \\
&\implies \alpha\gamma \frac{y}{l_b} = \alpha\gamma \frac{(N-n_b)^{1-\rho}}{(L-l_b)^{1-\alpha\gamma}} \left(\sum_{j\neq b} s_j\right)^{1-\gamma} \left(1 + \left(1 - \frac{\theta L}{l_b}\right)\varepsilon_{l_b r}^{-1}\right) + (1-\alpha)\gamma \frac{y}{l_b}\varepsilon_{l_b w}^{-1} \\
&\implies \gamma\left(s_b^{1-\gamma} n_b^{(1-\alpha)\gamma} l_b^{\alpha\gamma}\right)\left(\alpha - (1-\alpha)\varepsilon_{l_b w}^{-1}\right) = \alpha\gamma \frac{(N-n_b)^{1-\rho}}{(L-l_b)^{1-\alpha\gamma}} \left(\sum_{j\neq b} s_j\right)^{1-\gamma} \left(1 + \left(1 - \frac{\theta L}{l_b}\right)\varepsilon_{l_b r}^{-1}\right) l_b \\
&\implies \left[\frac{L-l_b}{l_b}\right]^{1-\alpha\gamma}\left(\alpha - (1-\alpha)\varepsilon_{l_b w}^{-1}\right) = \alpha \left(\frac{s_b}{\sum_{j\neq b} s_j}\right)^{\gamma-1} \left(1 + \left(1 - \frac{\theta L}{l_b}\right)\varepsilon_{l_b r}^{-1}\right)\left[\frac{N-n_b}{n_b}\right]^{(1-\alpha)\gamma}
\end{aligned}
$$

Replacing $(N-n_b)/n_b$ in the last line with the expression in equation (1.15) yields:

$$
\left[\frac{L-l_b}{l_b}\right]\left(1 - \frac{1-\alpha}{\alpha}\varepsilon_{l_b w}^{-1}\right)^{\frac{\rho}{1-\gamma}} = \left(\frac{\sum_{j\neq b} s_j}{s_b}\right)\left[1 + \left(1 - \frac{\theta L}{l_b}\right)\varepsilon_{l_b r}^{-1}\right]^{\frac{\rho}{1-\gamma}}, \tag{1.16}
$$

which implicitly defines the landlord's demand for land in terms of parameters only.

Proof that $l_b^*|_{\theta=0} < l_b^{pc}$:

Note that under perfect competition the landlord's optimal demand for land implies that:

$$
\begin{aligned}
l_b^{pc} &= L \times \frac{s_b}{\sum_j s_j} \\
\implies \frac{L}{l_b^{pc}} &= \frac{s_b + \sum_{j\neq b} s_j}{s_b} \\
\implies \frac{L - l_b^{pc}}{l_b^{pc}} &= \frac{\sum_{j\neq b} s_j}{s_b}.
\end{aligned}
$$

while the expression for the landlord's demand for land under market power shown in equation (1.16) when $\theta = 0$ is:

$$\left.\frac{L - l_b}{l_b}\right|_{\theta=0} = \left(\frac{\sum_{j\neq b} s_j}{s_b}\right)\left[\frac{\alpha(1+\varepsilon_{l_b r}^{-1})}{\alpha - (1-\alpha)\varepsilon_{l_b w}^{-1}}\right]^{\frac{\rho}{1-\gamma}},$$

Note that, based on the expressions for elasticities in equations (1.14):

$$\left[\frac{\alpha(1+\varepsilon_{l_b r}^{-1})}{\alpha - (1-\alpha)\varepsilon_{l_b w}^{-1}}\right] > 1 \iff \alpha(1-\alpha\gamma)\frac{l_b}{L - l_b} > (1-\alpha)\alpha\gamma\frac{l_b}{L - l_b}$$

$$\iff 1 > \gamma.$$

Which is always the case since $\gamma \in (0,1)$. The expressions for land demand above and the fact that $\left[\frac{\alpha(1+\varepsilon_{l_b r}^{-1})}{\alpha-(1-\alpha)\varepsilon_{l_b w}^{-1}}\right] > 1$ imply that:

$$\left.\frac{L - l_b}{l_b}\right|_{\theta=0} > \frac{L - l_b^{pc}}{l_b^{pc}} \implies l_b|_{\theta=0} < l_b^{pc}$$

■

Proof that $\frac{\partial l_b^*}{\partial \theta} > 0$:

From equation (1.16), define,

$$\Psi \equiv \left[\frac{L - l_b}{l_b}\right]^{\frac{1-\gamma}{\rho}}\left(1 - \frac{1-\alpha}{\alpha}\varepsilon_{l_b w}^{-1}\right) - \left(\frac{\sum_{j\neq b} s_j}{s_b}\right)^{\frac{1-\gamma}{\rho}}\left[1 + \left(1 - \frac{\theta L}{l_b}\right)\varepsilon_{l_b r}^{-1}\right] = 0.$$

By the implicit function theorem $\frac{\partial l_b^*}{\partial \theta} = -\frac{\partial\Psi/\partial\theta}{\partial\Psi/\partial l_b}$. Note that,

$$\frac{\partial \Psi}{\partial \theta} = \underbrace{\left(\frac{\sum_{j \neq b} s_j}{s_b} \right)^{\frac{1-\gamma}{\rho}}}_{(+)} \times \underbrace{\frac{\varepsilon_{l_b r}^{-1} L}{l_b}}_{(+)} > 0,$$

and so $\frac{\partial l_b^*}{\partial \theta} > 0 \iff \frac{\partial \Psi}{\partial l_b} < 0$.

To evaluate $\frac{\partial \Psi}{\partial l_b}$ it is convenient to define the term $\tilde{L} \equiv \left[\frac{L - l_b}{l_b} \right]^{\frac{1-\gamma}{\rho}}$, and to note that:

$$\frac{\partial \tilde{L}}{\partial l_b} \leq 0; \quad \frac{\partial \varepsilon_{l_b r}^{-1}}{\partial l_b} \geq 0; \quad \frac{\partial \varepsilon_{l_b w}^{-1}}{\partial l_b} \leq 0.$$

This derivative is then,

$$\frac{\partial \Psi}{\partial l_b} = \underbrace{\frac{\partial \tilde{L}}{\partial l_b}}_{(-)} \underbrace{\left(1 - \frac{1-\alpha}{\alpha} \varepsilon_{l_b w}^{-1} \right)}_{(+)} + \tilde{L} \underbrace{\left(\frac{1-\alpha}{\alpha} \frac{\partial \varepsilon_{l_b w}^{-1}}{\partial l_b} \right)}_{(-)} - \underbrace{\left(\frac{\sum_{j \neq b} s_j}{s_b} \right)^{\frac{1-\gamma}{\rho}}}_{(+)} \left[\underbrace{\frac{\theta L}{l_b^2} \varepsilon_{l_b r}^{-1}}_{(+)} + \left(1 - \frac{\theta L}{l_b} \right) \underbrace{\frac{\partial \varepsilon_{l_b r}^{-1}}{\partial l_b}}_{(+)} \right],$$

and thus when $\theta L < l_b$, it is always the case that $\frac{\partial \Psi}{\partial l_b} < 0$.

When $\theta L > l_b$, note that:

$$\begin{aligned} \frac{\partial \Psi}{\partial l_b} \leq 0 &\iff \frac{\theta L}{l_b^2} \varepsilon_{l_b r}^{-1} > \left(1 - \frac{\theta L}{l_b} \right) \frac{\partial \varepsilon_{l_b r}^{-1}}{\partial l_b} = \left(1 - \frac{\theta L}{l_b} \right) \frac{L}{(L - l_b) l_b} \varepsilon_{l_b r}^{-1} \\ &\iff \theta > \frac{l_b - \theta L}{L - l_b}, \end{aligned}$$

which always holds since $l_b - \theta L < 0$, $L - l_b \geq 0$, and $\theta \in (0, 1)$. ∎

Market power with land ceilings: expressions for prices, elasticities, and input demands

The Lagrangian associated to the maximization problem in equation (1.5) is:

$$\mathcal{L} = s_i^{1-\gamma}[(l_{Ui}+l_{Ri})^{\alpha} n_i^{1-\alpha}]^{\gamma} - r_U(l_{Ui}-l_{Ui}^0) - r_R(l_{Ri}-l_{Ri}^0) - w(n_i - n_i^0) \\ + \lambda(l_{Ri}-\bar{l}) + \mu_1 l_{Ui} + \mu_2 l_{Ri} + \mu_3 n_i,$$

with first-order conditions:

$$\frac{\partial \mathcal{L}}{\partial l_{Ui}} = \alpha\gamma s_i^{1-\gamma} n_i^{(1-\alpha)\gamma}(l_{Ui}+l_{Ri})^{(\alpha\gamma-1)} - r_U + \mu_1 = 0 \tag{1.17}$$

$$\frac{\partial \mathcal{L}}{\partial l_{Ri}} = \alpha\gamma s_i^{1-\gamma} n_i^{(1-\alpha)\gamma}(l_{Ui}+l_{Ri})^{(\alpha\gamma-1)} - r_R + \mu_2 = 0 \tag{1.18}$$

$$\frac{\partial \mathcal{L}}{\partial n_i} = (1-\alpha\gamma) s_i^{1-\gamma} n_i^{(1-\alpha)\gamma-1}(l_{Ui}+l_{Ri})^{(\alpha\gamma)} - w + \mu_3 = 0, \tag{1.19}$$

and complementary slackness conditions:

$$\lambda(l_{Ri}-\bar{l}) = 0; \quad \mu_1 l_{Ui} = 0; \quad \mu_2 l_{Ri} = 0; \quad \mu_3 n_i = 0.$$

Case *i):* $\lambda = 0$ (Unconstrained agent, $i \in \mathcal{S}$).

In this case $\mu_1 \neq 0$ and also, since the production function is strongly monotonically increasing, $\mu_2, \mu_3 = 0$. Then, from equation (1.18) and equation (1.19):

$$n_i^S = s_i \left(\gamma \left(\frac{\alpha}{r_R}\right)^{\alpha\gamma} \left(\frac{1-\alpha}{w}\right)^{1-\alpha\gamma}\right)^{\frac{1}{1-\gamma}}, l_{Ri}^S = s_i \left(\gamma \left(\frac{\alpha}{r_R}\right)^{\rho} \left(\frac{1-\alpha}{w}\right)^{1-\rho}\right)^{\frac{1}{1-\gamma}},$$

and $l_{Ui}^S = 0$.

Case *ii):* $\lambda > 0$ and $\mu_1 > 0$ (Mid-ability constrained agent, $i \in \mathcal{C}_1$).

In this case both $l_{Ri}^{C_1} = \bar{l}$ and $l_{Ui}^{C_1} = 0$. Hence, from equation (1.19) the demand for labor for this type of agent is:

$$n_i^{C_1} = s_i^{\frac{1-\gamma}{\rho}} \left(\frac{(1-\rho)\bar{l}^{\alpha\gamma}}{w} \right)^{\frac{1}{\rho}},$$

Case *iii):* $\lambda > 0$ and $\mu_1 = 0$ (High-ability constrained agent, $i \in \mathcal{C}_2$).

In this case $l_{Ri}^{C_1} = \bar{l}$, and from equations (1.17) and (1.19):

$$n_i^{C_2} = s_i \left(\gamma \left(\frac{\alpha}{r_U} \right)^{\alpha\gamma} \left(\frac{1-\alpha}{w} \right)^{1-\alpha\gamma} \right)^{\frac{1}{1-\gamma}}, \quad l_{Ui}^{C_2} = s_i \left(\gamma \left(\frac{\alpha}{r_U} \right)^{\rho} \left(\frac{1-\alpha}{w} \right)^{1-\rho} \right)^{\frac{1}{1-\gamma}} - \bar{l}.$$

Now define for conveniency the following auxiliary terms:

$$\mathcal{I}_{C_1} \equiv \sum_{i \in C_1} \bar{l}; \quad \mathcal{I}_{C_2} \equiv \sum_{i \in C_2} \bar{l}; \quad \Sigma_{C_1} \equiv \sum_{i \in C_1} s_i; \quad \Sigma_{C_2} \equiv \sum_{i \in C_2} s_i; \quad \Sigma_S \equiv \sum_{i \in S} s_i$$
$$\mathcal{L}_R \equiv \psi(1-\theta)L - \mathcal{I}_{C_1} - \mathcal{I}_{C_2} - \bar{l}; \quad \mathcal{L}_U \equiv (1 - \psi + \psi\theta)L - l_{Ub} + \mathcal{I}_{C_2}.$$

Combining the individual input demand functions shown above with the market clearing conditions yields the following expressions:

$$
\begin{aligned}
L_R = \psi(1-\theta)L &= \sum_{i\in S} l_{Ri}^{S} + \sum_{i\in C_1} l_{Ri}^{C_1} + \sum_{i\in C_2} l_{Ri}^{C_2} + l_{Rb} = \sum_{i\in S} l_{Ri}^{S} + \underbrace{\sum_{i\in C_1}}_{=\mathcal{I}_{C_1}} \bar{l} + \underbrace{\sum_{i\in C_2}}_{=\mathcal{I}_{C_2}} \bar{l} \\
\implies \mathcal{L}_R &= \sum_{i\in S} s_i \left(\gamma \left(\frac{\alpha}{r_R}\right)^{\rho} \left(\frac{1-\alpha}{w}\right)^{1-\rho}\right)^{\frac{1}{1-\gamma}}, \qquad (1.20)
\end{aligned}
$$

$$
\begin{aligned}
L_U = (1-\psi+\psi\theta)L &= \sum_{i\in S} l_{Ui}^{S} + \sum_{i\in C_1} l_{Ui}^{C_1} + \sum_{i\in C_2} l_{Ui}^{C_2} + l_{Ub} \\
\implies \mathcal{L}_U &= \sum_{i\in C_2} s_i \left(\gamma \left(\frac{\alpha}{r_U}\right)^{\rho} \left(\frac{1-\alpha}{w}\right)^{1-\rho}\right)^{\frac{1}{1-\gamma}}, \qquad (1.21)
\end{aligned}
$$

$$
\begin{aligned}
N &= \sum_{i\in S} n_i^{S} + \sum_{i\in C_1} n_i^{C_1} + \sum_{i\in C_2} n_i^{C_2} + n_b \\
&= \left(\gamma\alpha^{\alpha\gamma}\left(\frac{1-\alpha}{w}\right)^{1-\alpha\gamma}\right)^{\frac{1}{1-\gamma}} \left(\frac{\sum\limits_{i\in S} s_i}{r_R^{\frac{\alpha\gamma}{1-\gamma}}} + \frac{\sum\limits_{i\in C_2} s_i}{r_U^{\frac{\alpha\gamma}{1-\gamma}}}\right) + \sum_{i\in C_1} s_i^{\frac{1-\gamma}{\rho}} \left(\frac{(1-\rho)\bar{l}^{\alpha\gamma}}{w}\right)^{\frac{1}{\rho}} + n_b. \qquad (1.22)
\end{aligned}
$$

Define the additional auxiliary term:

$$
\Phi \equiv \left(\mathcal{L}_R^{\alpha\gamma}\Sigma_S^{1-\gamma}\right)^{\frac{1}{\rho}} + \bar{l}^{\frac{\alpha\gamma}{\rho}} \sum_{i\in C_1} s_i^{\frac{1-\gamma}{\rho}} + \left(\mathcal{L}_U^{\alpha\gamma}\Sigma_{C_2}^{1-\gamma}\right)^{\frac{1}{\rho}}, \qquad (1.23)
$$

combining equations (1.20), (1.21), and (1.22) yields an expression for the wage in terms of parameters and landlords' choices only:

$$w = \frac{(1-\alpha)\gamma}{(N-n_b)^\rho} \left[\left(\underbrace{\psi(1-\theta)L - \mathcal{I}_{C_1} - \mathcal{I}_{C_2} - \bar{l}}_{\equiv \mathcal{L}_R} \right)^{\frac{\alpha\gamma}{\rho}} \left(\underbrace{\sum_{i \in S} s_i}_{\equiv \Sigma_S} \right)^{\frac{1-\gamma}{\rho}} + \bar{l}^{\frac{\alpha\gamma}{\rho}} \sum_{i \in C_1} s_i^{\frac{1-\gamma}{\rho}} \right.$$

$$\left. + \left(\underbrace{(1-\psi+\psi\theta)L - l_{Ub} + \mathcal{I}_{C_2}}_{\equiv \mathcal{L}_U} \right)^{\frac{\gamma\alpha}{\rho}} \left(\underbrace{\sum_{i \in C_2} s_i}_{\equiv \Sigma_{C_2}} \right)^{\frac{1-\gamma}{\rho}} \right]^\rho = \frac{(1-\alpha)\gamma}{(N-n_b)^\rho} \Phi^\rho,$$

and, analogously, for land prices:

$$r_U = \frac{\alpha\gamma(N-n_b)^{1-\rho}}{\Phi^{1-\rho}} \times \left(\frac{\Sigma_{C_2}}{\mathcal{L}_U} \right)^{\frac{1-\gamma}{\rho}}$$
$$r_R = \frac{\alpha\gamma(N-n_b)^{1-\rho}}{\Phi^{1-\rho}} \times \left(\frac{\Sigma_S}{\mathcal{L}_R} \right)^{\frac{1-\gamma}{\rho}}.$$

Finally, the corresponding inverse demand elasticities are then:

$$\varepsilon_{l_b, r_U}^{-1} \equiv \frac{\partial \ln r_U}{\partial \ln l_b} = \frac{(1-\rho)\alpha\gamma}{\rho} \times \frac{l_b}{\Phi} \left(\frac{\Sigma_{C_2}}{\mathcal{L}_U} \right)^{\frac{1-\gamma}{\rho}} + \frac{1-\gamma}{\rho} \left(\frac{l_b}{\mathcal{L}_U} \right)$$

$$\varepsilon_{l_b, r_R}^{-1} \equiv \frac{\partial \ln r_R}{\partial \ln l_b} = \frac{(1-\rho)\alpha\gamma}{\rho} \times \frac{l_b}{\Phi} \left(\frac{\Sigma_{C_2}}{\mathcal{L}_U} \right)^{\frac{1-\gamma}{\rho}}$$

$$\varepsilon_{l_b, w}^{-1} \equiv \frac{\partial \ln w}{\partial \ln l_b} = -\alpha\gamma \times l_b \left(\frac{\Sigma_{C_2}}{\mathcal{L}_U} \right)^{\frac{1-\gamma}{\rho}}.$$

1.10 Appendix C: Data Appendix

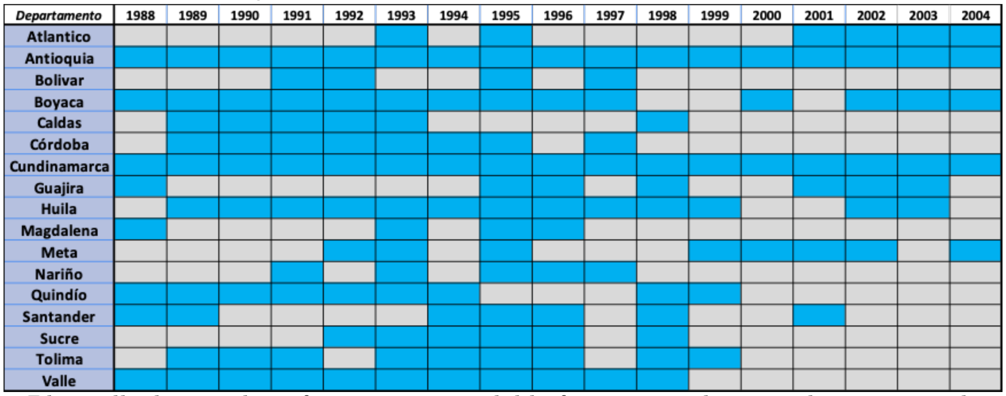
Municipal agricultural productivity - *Evaluaciones Agropecuarias Municipales*

The municipal agricultural evaluations were first established in 1972 by the ministry of agriculture for the purpose of obtaining biannual information on agricultural production. Since their inception they were conceived as a way to process, aggregate, and harmonize information coming from various state and private actors linked to the agricultural sector. Hence their informal name of agricultural consensus (*consensos agrícolas*). The process of collecting each agricultural evaluation entailed the distribution of surveys enquiring about the quantities planted, harvested, and produced for a wide range of agricultural products during the course of the previous harvesting season. The intended respondents of the surveys were both large and small agricultural producers and cooperatives, extension officers, agrochemical input distributors, community leaders, and local community councils. After collecting the responses, government officials were required to schedule a meeting to present their preliminary findings, and to conduct field visits when there were large enough disagreements in the figures presented.

While the methodological guidelines to collect the evaluations were defined at the national level by the ministry of agriculture, between 1980 and 2006, the execution of the evaluations—in addition to the processing of the data and the publishing of the results—relied on *departamento*-level government dependencies known as the regional units for agricultural planning (*Unidades regionales de planeación agrícola*) (URPA).[23] Starting in 1980, the agricultural evaluations began presenting information disaggregated at the municipal level. Every *departamento*-level agency would usually publish two reports per year (semesters A and B) with information for transitory, perennial and annual crops in all municipalities under their jurisdiction. While the amount of detail in each publication varies by agency and year, all of them include crop-specific information on area planted, area harvested, and quantity produced, with most reports also including information on production costs, and farm gate prices, or even crop quality. Starting in 2007 local government offices stopped being responsible for the publication of these reports, and the *Evaluaciones* started to be published centrally by the national ministry of agriculture. For this project I collected and digitized hundreds of these half-yearly reports for the period 1988–2004, but, as described in Section 3.2 in the body of the text, this data collection process is still a work in progress. Figure 1.13 illustrates the *Evaluaciones* collected for this project, and Figures 1.14, and 1.15 show examples of the usual data layout within these publications.

[23] *Departamentos* are the second-level sub-national administrative divisions in Colombia, akin to U.S. states. *Departamentos* are collections of municipalities (*municipios*), which are analogous to U.S. counties.

Figure 1.13: *Evaluaciones Agropecuarias Municipales* - List of Publications Collected and Digitized



Notes: Blue cells denote the information is available for municipalities in that state in that year.

Figure 1.14: Agricultural Evaluations - Examples of Publication Covers

Figure 1.15: Agricultural Evaluations - Example of Data Table Layouts

Cuadro No 2:Cundinamarca.Cultivos Permanentes por Municipios
Evaluacion Municipal. Mayo de 1991

Crop type (Perennial)
Year

Comite Desarrollo Agropecuario-Secretaria Agricultura-URPA

Municipio	Cultivo	Area Plantada Eval.90 (ha)	Area en Producc. Eval.90 (ha)	Area Plantada Prog. 91 (ha)	Area Produc. Prog.91 (ha)	Produc. Eval. 1990 (ton)	Produc. Program. 1991 (ton)	Rendim. Eval. 1990 (kg/ha)	Rendim. Progra. 1991 (kg/ha)
		Area Planted	Area Harvested			Production		Yield	
Municipality									
** Municipio: Alban									
	Platano Crop	70.0	70.0	70.0	70.0	490.0	490.0	7000.0	7000.0
** Subtotal **									
		70.0	70.0	70.0	70.0	490.0	490.0		
Municipality									
** Municipio: Anapoima									
	Cana Panela	130.0	130.0	150.0	130.0	780.0	780.0	6000.0	6000.0
	Citricos	260.0	260.0	280.0	260.0	1300.0	1300.0	5000.0	5000.0
	Mango Crops	900.0	550.0	900.0	600.0	3025.0	3300.0	5500.0	5500.0
	Maracuya	380.0	380.0	430.0	400.0	2660.0	2800.0	7000.0	7000.0
** Subtotal **									
		1670.0	1320.0	1760.0	1390.0	7765.0	8180.0		
** Municipio: Anolaima									
	Cacao	50.0	40.0	60.0	50.0	20.0	25.0	500.0	500.0
	Citricos	500.0	495.0	520.0	500.0	14850.0	15000.0	30000.0	30000.0
	Mora	30.0	30.0	30.0	30.0	240.0	240.0	8000.0	8000.0
	Platano	800.0	300.0	500.0	340.0	4500.0	2550.0	15000.0	7500.0
** Subtotal **									
		1380.0	865.0	1110.0	920.0	19610.0	17815.0		

(a) Cundinamarca - 1990

MINISTERIO DE AGRICULTURA Y DESARROLLO RURAL - SECRETARIA DE FOMENTO

Crop Type (Anual) | Year

CULTIVOS ANUALES 1 EVALUACION DEFINITIVA DEL AÑO DE 1990

Cultivo	Municipio	Area cos. Ha.	Producción Ton.	Rendimiento Kg/Ha.
Crop	Municipality	Area Harvested	Production	
ALCACHOFA				
	TOCA	2	3	1500
Subtotal	1	2	3	1500
ARRACACHA				
	ALMEIDA	7	49	7000
	BOYACA	488	2928	6000
	CIENEGA	153	1071	7000
	GARAGOA			
	GUATEQUE	30	240	8000
	GUAYATA	10	70	7000
	JENESANO	150	900	6000
	LA CAPILLA	20	140	7000
	MONIQUIRA			
	SOMONDOCO	7	49	7000
	SUTATENZA	15	120	8000
	TENZA	45	360	8000
	TIBANA	75	495	6600
	TOGUI	30	240	8000
	UMBITA	35	245	7000
	VIRACACHA	28	196	7000
Subtotal	16	1093	7103	6499
CEBOLLA JUNCA				
	AQUITANIA	1800	108000	60000
	TOTA	5	135	27000
Subtotal	2	1805	108135	59909
HABA				
	AQUITANIA	50	58	1150

(b) Boyacá - 1990

Chapter 2

Temperature Shocks and Land Fragmentation: Evidence from Transaction and Property Registry Data

With Nicolás de Roux, Margarita Gáfaro, Ana María Ibáñez, and Heitor Pellegrina.

2.1 Introduction

Large shares of the population in low and middle-income countries are employed in small, low-productivity farms (Restuccia et al., 2008; Adamopoulos and Restuccia, 2014; Gollin et al., 2014). The prevalence of small farms can constrain technological progress and limit potential economies of scale and productivity gains, hindering poverty reduction and development (Foster and Rosenzweig, 2022). Understanding the determinants of the farm size distribution is therefore a first-order concern.

This paper studies a potential determinant of land fragmentation that is particularly salient in low and middle-income countries, uninsured risk. In these settings, agricultural production is highly exposed to income fluctuations related to weather and commodity price variability and coping mechanisms like insurance and credit are scant (Jayachandran, 2006; Colmer, 2021; Fafchamps, 1992; Cole et al., 2017; Carter et al., 2017).[1] In the event of a negative productivity shock, poor land owners may have to sell a fraction of their landholdings in order to smooth consumption, affecting in turn the farm-size distribution (Rosenzweig and Wolpin, 1993; Carter and Zimmerman, 2003; Kazianga and Udry, 2006).[2] Using two unique administrative data sets with information on hundreds of thousands of land sale transactions and information on a land registry covering most of the country of Colombia, we show that temperature shocks cause land sales and lead on average to smaller-sized farms. We show that this reduction in farm size is entirely driven by the entry of new landholders that operate relatively small farms, and find no evidence of shocks leading to the consolidation of larger holdings. To explain these patterns we develop a general equilibrium, heterogeneous-agent model where agents face an intertemporal consumption decision bound to a subsistence constraint. The model illustrates how the occurrence of negative productivity shocks can lead to the exit of incumbent farmers from the agricultural sector, while also causing new, previously landless agents to buy land. Both our empirical and theoretical results document how climate-induced distress land sales are a relevant margin of adjustment that can have important distributional and productivity implications. The results shed light on an additional, potentially large, negative consequence of climate change, given that the intensity and frequency of extreme weather events are bound to increase in the coming decades (IPCC, 2021).

To study the relationship between extreme weather events, land transactions and farm size, we use a unique administrative dataset containing official records of land transactions between 2000 and 2011 involving plots allocated by the Colombian government to private farmers throughout the 20$^{\text{th}}$ century. These plots comprise about 50% of all rural land currently held by private individuals in the country and are evenly distributed across regions. With information on nearly 500,000 land transactions we construct a yearly balanced panel with the number of full and partial land sales both at the municipality level and at the *vereda* level, Colombia's smallest rural administrative unit. We complement this data with information collected from the National Land Registry, a census of

[1] Unsubsidized agricultural insurance coverage rates in high income countries are on average 41.7% while coverage rates for lower-middle income and low income countries are, respectively, 8% and 0.5% (Mahul and Stutley, 2008).

[2] According to a longitudinal survey of rural Colombian households, between 2013 and 2016, nearly 65% of households who reported selling land did so in order to pay for household expenses or cover outstanding debts, pay for a medical treatment, or pay for education fees. These figures come from the ELCA survey described in more detail below.

properties covering most of Colombia's farmland. This dataset allows us to measure yearly changes in the number of land owners and the distribution of plot sizes at the municipality level. Because land rental markets in Colombia are thin –data from a national representative survey of farms shows that in 2019 only 9% of farms operated rented land–,[3] these measures of plot size are a good representation of farm size and farm operational scale.

We combine both datasets with high-resolution meteorological data from Copernicus Climate Change Service (C3S). Our preferred measure of temperature shocks identifies days of atypically high or low temperatures by constructing distributions that are specific to the vereda (or the municipality) and to the calendar quarter. This accounts for seasonality and for differences across regions in weather patterns. We exploit both within-vereda and within-municipality variation in weather shocks to identify the causal effects of interest under the standard assumption in the literature (e.g., Dell et al. (2014)) that, conditional on time and geographical unit fixed effects, temperature shocks are uncorrelated with other time-varying factors affecting land sales.

First, we show that extreme temperature shocks induce distress land sales. In particular, 100 additional days of atypical temperature in a two-year period increase the number of land sales in the municipality by 7.6%. These temperature shocks also induce land fragmentation as average farm size decreases by 1.2%. The latter is driven by the entry of new owners with land holdings in the lowest quintiles of the initial size distribution. The effect of weather shocks on land sales is stronger in less densely populated municipalities, located farther away from urban markets. While land owners in wealthier, better connected municipalities are more likely to respond to negative temperature shocks by taking out mortgages on their land. This suggest that better access to credit can mitigate the need for distress sales. We complement our main findings using data from a 3-wave longitudinal household survey and show that following an adverse temperature shock, rural households have lower consumption, are more likely to migrate, are less likely to hold land, and are more likely to reallocate their labor to the non-agricultural sector. These effects are consistent with the use of distress sales as a consumption smoothing mechanism.

This paper contributes to the literature that explores the determinants of farm size in developing

[3] National Agricultural Survey (ENA), carried out by the National Statistical Agency (DANE); 2019-1 bulletin.

countries. Recent literature on this topic has focused on institutional factors that distort farm sizes and induce misallocation (Adamopoulos and Restuccia, 2020; Chen et al., 2022), or on the changes to the distribution of farm sizes induced by variations in urban labor demand (Rao et al., 2022; Madhok et al., 2022). We add to this literature by providing evidence on the effect of negative productivity shocks on farm size. While a standard heterogeneous-agent model with credit market imperfections would predict that the expansion in land supply due to distress sales should lead to the consolidation of small farms into larger landholdings, we show that the opposite effect, land *fragmentation*, takes place.

Our results also emphasize that low agricultural productivity can be exacerbated by the aggregate consequences of individual responses to uninsured risk. By documenting how the aggregate exposure to adverse weather shocks leads to a more fragmented farm size distribution, our findings point to another mechanism explaining the notoriously low productivity of agriculture relative to the non-agricultural sector in developing economies (Gollin et al., 2014; Restuccia et al., 2008; Caselli, 2005). While some previous studies have documented the occurrence of distress land sales with survey data in several developing countries (Cain, 1981; Deininger and Jin, 2008; Musyoka et al., 2021), our use of administrative data allows us to estimate the aggregate effects of distress sales on the farm size distribution.

Finally, this paper contributes to the literature exploring the effects of weather shocks on agriculture. This literature has shown that farmers' responses to weather shocks include adjustments in labor and intermediate inputs use, changes in crop choice, migration, or investment in human capital (Jayachandran, 2006; Jessoe et al., 2018; Colmer, 2021; Jagnani et al., 2021; Aragón et al., 2021). We complement this literature by documenting that land sales constitute an important margin of adjustment for farmers facing negative productivity shocks. Because land is the main financial asset of most farmers in developing economies, land sales can have strong, long-lasting effects on farmers' future income. As climate change intensifies, our results highlight an additional mechanism through which increases in the severity and frequency of adverse weather shocks can deepen the wedge in the performance of agricultural sectors between poor and rich economies (Burke et al., 2015; IPCC, 2021).

2.2 Context, Data, and Descriptive Statistics

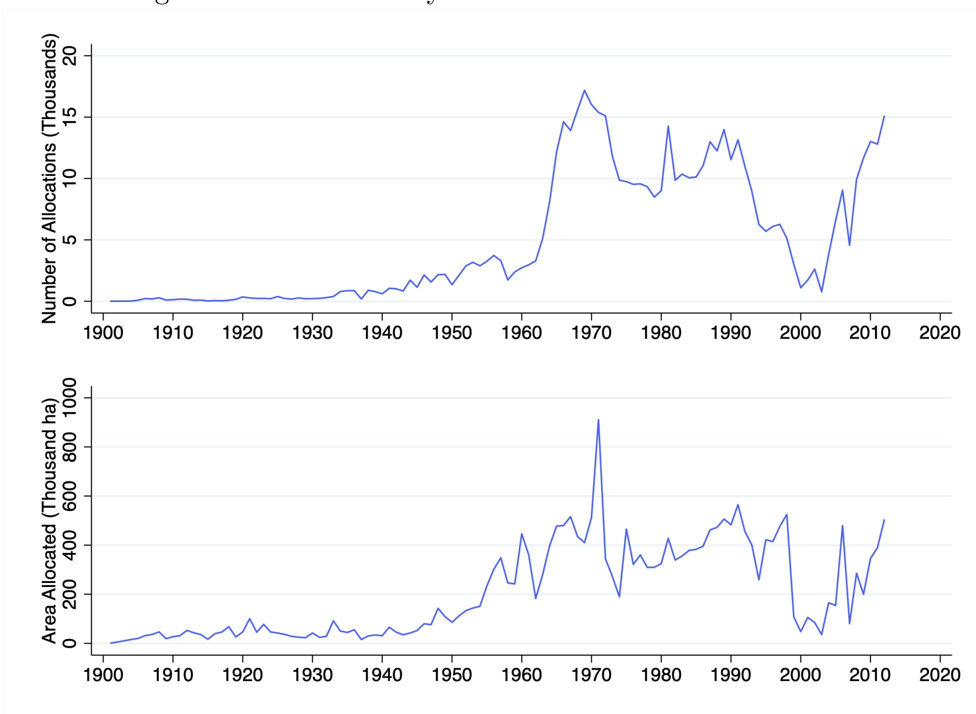
Studying the relation between land market transactions, land fragmentation, and weather shocks requires information with special characteristics. First, we need information at the transaction level spanning a long time period and a large geographical area. Second, assessing land fragmentation requires a registry of plot information that allows for characterizing the complete distribution of the size of farms for a given geographical unit. Third, we need measures of weather shocks that are homogeneous across time and space and that can be linked to the transaction and land registry data at some fine geographical level. In this paper, we use two unique administrative data sets that allow us to study the relations of interest at an extremely granular level. The first one contains information on plots that were originally granted to owners in the context of the Colombian public land distribution program. The second contains information from land property registries. In this section, we provide an account of the institutional and historical context associated with land redistribution in Colombia and describe the different data sets that we use.

2.2.1 The public land allocation program and the transaction data

The Colombian government has carried out free allocations of public idle lands (*baldíos*) to private individuals uninterruptedly since the beginning of the twentieth century. This allocations have become the largest and most consequential land reform policy instrument employed by the national government (Albertus, 2015). Formally, a *baldío* allocation is an administrative resolution issued by the national government to transfer state-owned vacant land to a private party. This allocation process has mostly consisted of a combination of frontier-settlement schemes where unused public lands are granted to poor smallholders, and of programs focused on the titling of state-owned lands that might have been previously informally occupied (Ibáñez and Muñoz, 2010).

The bulk of government-owned land allocations began in the midst of the US *Alliance for Progress* program with the enactment of the Social Agrarian Reform Act (Law 135) in 1961, which established the land reform agency (INCORA, later renamed as INCODER, and currently the National Land Agency, ANT). During the second half of the twentieth century, land allocation laws were amended on three occasions (Law 01 of 1968, Law 30 of 1988, and Law 160 of 1994) but the explicit objective of the policy always remained that of reducing land inequality and giving land to landless farmers (CNMH, 2016). Figure 2.1 shows the evolution of baldíos allocations since 1901, the vast majority of which were granted between 1960 and 1990. In terms of the number of beneficiaries and the amount of land allocated, the scale of the policy has been vast. More than 550,000 land plots have been granted to private individuals in 1,034 of the 1,122 existing municipalities. These plots account for 23 million hectares –more than half of the currently privately-held land in the country (Sánchez and Villaveces, 2016; Arteaga et al., 2017).

Figure 2.1: One Century of Land Allocations - 1901–2012



Notes: Data from the System of Information for Rural Development (SIDER)

Land petitioners undergo an administrative process with the national land agency to determine if they fulfill the legal requirements to become a beneficiary. While the requirements have changed in time, the most important conditions petitioners must fulfill involve owning no other land and having an income below a given threshold. Under the current legislation, the process formally consists of nine steps, which include the placement of an ad announcing the allocation in a local newspaper, and a physical inspection of the plot to be granted. Although on paper this procedure should take 60 days, allocation processes are generally much lengthier and some can take years

(Gutiérrez Sanín, 2019). Appendix Figure 2.6 shows the evolution of the average and median size of allocated plots since 1960. The overwhelming majority of land allocations made throughout 1961–2014 period consisted of relatively small land plots, with a median allocation size across municipalities of 6.6 hectares. Importantly for this paper, Law 160 of 1994 established a ceiling on the amount of government-allocated land to which a single individual can claim ownership. This limit, defined by the municipality-specific Agricultural Family Unit (UAF), restricts the capacity of relatively larger farmers to purchase land that was initially government-owned. In appendix section 2.8, we show that these land ceilings are not driving our results.

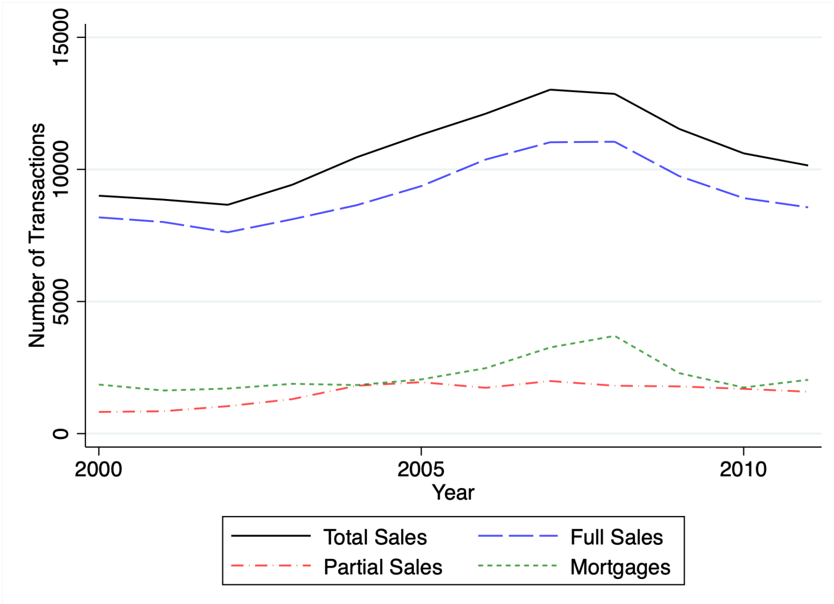
The universe of land allocations made by the government throughout 1901–2011 period is registered in the System of Information for Rural Development (SIDER) dataset currently maintained by the ANT. After receiving the plot, beneficiaries must register the property in the office of the local public notary, and all formal land transactions carried out over the plot (including mortgages) are henceforth registered and stored in a dataset maintained by the National Superintendence of Notaries (SNR), the government agency that supervises regional notaries and keeps a record of all real estate market transactions held among private parties.[4]

Our main source of data is the transaction history of all baldío allocations whose beneficiaries registered their property with the notary thus finalizing the process to obtain a formal property right.[5] We mainly focus on land purchase transactions, which can be either the transfer of an entire property from one individual to another, or the subdivision and sale of only a fraction of the original plot. We refer to these types of transactions as *full sales* or *partial sales* respectively. We also study mortgages, as they could constitute an important adjustment margin when coping with negative productivity shocks. For each transaction held between two parties, we have access to information on the plot's location, the date in which it occurred, and the type of transaction. Figure 2.2 shows the yearly evolution of full and partial sales, along with the number of mortgages

[4]The history of the transactions carried out over a plot, named the Certificate of Liberty and Tradition (*Certificado de Libertad y Tradición*) is public information that can be consulted by paying a small fee for any property with a real estate registration number on the web page of the SNR.

[5]While the registration process was not automatic and a non-negligible number of beneficiaries failed to follow this last administrative step (Faguet et al., 2020), Appendix Figure 2.7 in the appendix shows that allocations and real estate registrations follow each other closely across time, suggesting that the great majority of land plots allocated did end up being registered.

Figure 2.2: Yearly land transactions - 2000–2011



Notes: Data from the National Superintendence of Notaries (SNR). The figure shows the national-level yearly number of transactions held over plots originally granted by the national government.

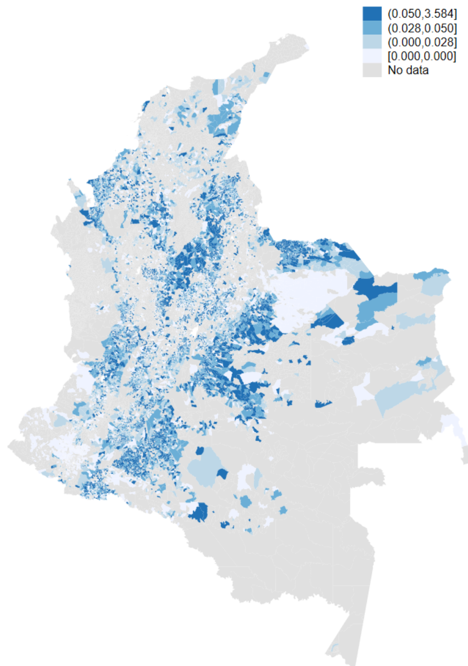
originated. Most of the sales in the land market are full sales, with partial sales representing a relatively small fraction of total transfers.

We match the location of the plot in the SNR dataset to the official list of Colombian municipalities and veredas provided by DANE, Colombia's National Statistical Agency.[6] We construct a balanced yearly panel both at the municipality and at the vereda level with information on the number of full and partial land sales, mortgages, and government land allocations. While we can match each of the land plots in the SNR data to their corresponding municipality, not all properties have information on the vereda, and we are able to identify it for only 63% of the properties in the SNR data. Figure 2.3 shows the ratio of total land sales to total allocations for the sample of plots matched to a vereda between 1980 and 2010. The map shows that there is substantial variation in the amount of land sales across space and in the veredas for which we observe transactions.

When deciding on the adequate level of data aggregation we face a tradeoff between the coarser

[6]Municipalities are the smallest official administrative division in Colombia. For some administrative purposes, rural areas within municipalities are further divided into veredas. Veredas operate under the executive power of municipalities' mayors but have their own democratically elected Community Action Boards (*Juntas de Acción Comunal)*. There are approximately 30,000 veredas in Colombia and 1,123 municipalities.

Figure 2.3: Ratio of Land Sales to Number of Allocations



Notes: Data from the National Superintendence of Notaries (SNR). The figure shows the proportion of plots sold in each vereda to the total number of plots allocated by the government between 1980 and 2011.

municipality level and the finer, but potentially selected, vereda sample. We estimate the effects of weather shocks on land transactions using both samples and present the results in Section 3.4. Reassuringly, the choice of sample does not affect the sign or statistical significance of the results.

2.2.2 The land registry

For over 50 years, the National Geographical Institute of Colombia (IGAC) has collected information on land use and ownership and keep land valuations up to date. Law 14 of 1983, instituted a plot-level information collection system (the 'Ficha Predial' system) which has been implemented and maintained by IGAC since then. This system is meant to collect information on the location, size, and economic purpose of all real properties in every Colombian municipality with the exception of the state of Antioquia, which runs its own, independent, cadastral information system (Ibánez et al., 2012).

This information system is meant to be an up-to-date census of land ownership for the whole country, and the law stipulates that IGAC must carry out cadastral updates in every municipality every five years. Information is not, however, updated on a regular basis and the amount of time between cadastral updates varies significantly across municipalities.[7] Martinez (2019) shows that IGAC updates are not driven by changes in economic conditions of the municipalities (e.g. property booms).

In our study, we use municipal-level aggregate information from all plots in IGAC's cadastre that are i) privately owned, and ii) categorized as having an agricultural economic purpose. This amounts to roughly 40 million hectares of land. We use a yearly panel of municipalities with the number of plots, the number of owners and average plot size within size ranges as calculated by (Ibánez et al., 2012). The data from the land registry is only available for the period 2000-2011 and so we restrict our analysis to this time period. We exclude from our final sample of municipalities (both for the transaction-level data and for the land registry data) large metropolitan areas and municipalities with very few (i.e. below the 99^{th} percentile) properties registered. Our final sample is made up of 927 municipalities, which encompass 85.3% of the rural population in the country.

2.2.3 Weather data and temperature shocks

We define temperature shocks that are specific to each geographical unit (either municipality or vereda) in order to account for the very large variation in climatic conditions across Colombian rural areas. The shocks are defined based on the unit's specific distribution of weather realizations, which we compute using long-run daily weather measurements (similar, for example, to Kaur (2019)). While this approach contrasts with weather shock definitions based on a fixed temperature threshold, which might be more suitable for the analysis of a specific region or crop (see, for example, Ibáñez et al. (2022)), we show that our results are robust to measures such measures of shocks that use using fixed thresholds.

We construct measures of temperature shocks using the ERA5 data set, provided by the Copernicus Climate Change Service (C3S) of the European Centre for Medium-Range Weather

[7] There are currently 80 municipalities across the country in which IGAC has not yet established the census-level cadastral information system. These municipalities have, instead, a self-reported information system ('Catastros Fiscales') in which landowners voluntarily register their properties in regional IGAC offices.

Forecasts (ECMWF). This dataset contains global reanalysis information on temperature with a horizontal resolution of 0.25×0.25 degrees (approximately 28 km^2 depending on the longitude) at an hourly frequency.[8] We use the temperature of the atmosphere two meters above the surface (in degrees Kelvin) from 1979 to 2016 in ERA5 for pixels in mainland Colombia. For each pixel in the data, we compute the average temperature for each day d, and obtain the average daily temperature of each vereda-day (or municipality-day) pair (v, d) by taking a weighted average of the pixels in the vereda using as weights the area of the pixel relative to the total area of the vereda. We compute the historical quarterly distribution of daily temperatures by considering all temperature measurements for pairs v, d in calendar-quarter q throughout the period 1979–2016. For each vereda this results in four distributions, one per quarter. We compute the 20$^{\text{th}}$ and 80$^{\text{th}}$ percentiles of each distribution and define the average temperature of a given vereda-day as atypically high if it is above the 80$^{\text{th}}$ percentile of the corresponding distribution of average daily temperatures of v, q. Analogously, we define a day as having atypically low temperatures if it is below the 20$^{\text{th}}$ percentile of the corresponding distribution.

Finally, for each year y, we sum the number of atypically high or low temperature days in each quarter. In our baseline specifications, we estimate the effect on outcomes measured at the vereda-year (v, y) frequency and use as our preferred measure of weather shock the total number of days with atypical temperatures over the past two years (i.e. $y - 1$, $y - 2$). Figures 2.8 and 2.9 in the appendix show the spatial and temporal variation of the resulting temperature shock measures across veredas. This definition of temperature shocks has two advantages. First, it takes into account seasonality at the calendar quarter level since the distribution is specific to q. For example, since some calendar quarter of the year are typically hotter, we only consider a day as atypically hot if the temperature is high relative to the historical temperature of that quarter. Second, the measure is specific to the vereda (or municipality) and takes into account that an absolute temperature might be atypically high and have a negative consequence in one place but not in another.

In the empirical exercises below we also control for total rainfall. To construct this measure we use the ERA5 monthly precipitation reanalysis data with resolution 0.1×0.1 degrees (approximately

[8] Reanalysis weather information from the ERA5 results from the combination of climate models and observational data from satellites and ground sensors.

9 km^2 depending on the longitude) and use the conversion factor provided C3S to obtain a measure of total monthly precipitations in cubic milliliters for each pixel. We then obtain a weighted average across the pixels in the vereda to obtain monthly average rainfall. Again, we use as weights the size of the pixel relative to the size of the vereda. For a given year, we add across months to obtain a measure of total precipitation in the pair vereda-year v, y. We take an analogous average of the pixels that compose a municipality to obtain measures of total yearly rainfall in a municipality.

Linking the weather data with the SNR land sales vereda-level panel yields a data set with 12,472 veredas across 782 municipalities. Panel A of Table 2.1 shows descriptive statistics of this sample. In a given vereda year, there are, on average, 18 accumulated adjudications, 0.55 sales –0.47 full sales and 0.07 partial sales–, and 0.11 mortgages. These numbers are low but there is considerable variation across veredas. On average there are 281 days of atypical temperature days in the two previous years. Linking weather data to the panel of yearly sales at the municipality level (Panel B of Table 2.1) yields a sample of 866 municipalities. On average there are 12.3 land sales on each municipality-year (10.6 full sales; 1.8 partial), and 2.6 mortgage originations. The average municipality-year observation had 277.2 days with atypical temperatures during the two past years, with a standard deviation of 56.3 days.

Finally, linking the temperature shock measures with the land registry panel yields a sample of 927 municipalities. In the average municipality-year, there are 2516 owners, 2519 farms, the size of the average farm is 29.4 hectares, and there were 277 days of atypical temperature in the past two years. Data in all samples is restricted to the 2000–2011 period.

2.2.4 Longitudinal household survey and additional data sources

We complement the previous data sources with data from a household panel that we use to analyze how farmers' decisions change in response to temperature shocks. In particular, we use the Colombian Longitudinal Survey conducted by the Universidad de los Andes (ELCA). The ELCA includes a sample of 4,800 rural households interviewed over three survey rounds (a baseline collected in 2010 and two follow-ups in 2013 and 2016). The rural sample of the ELCA is representative of small agricultural producers in four micro-regions: Atlantic, Central, Coffee-

Table 2.1: Descriptive Statistics

	Mean	Std. Dev.	Min	Max
	Panel A: SNR - Vereda (N = 12,472)			
Total number of sales	0.55	2.07	0	133
Number of full sales	0.47	1.80	0	132
Number of partial sales	0.07	0.64	0	61
Number of Mortgages	0.11	0.56	0	29
Days of atypical temperature	281.38	55.18	96	560
Days of atypical high temperature	158.42	93.46	0	508
Days of atypical low temperature	122.96	87.65	4	560
Number of total allocations	18.56	55.36	0	2,376
Accumulated precipitation	3,272.2	2,370.8	374.6	33,533
	Panel B: SNR - Municipality (N = 866)			
Total number of sales	12.38	24.56	0	292
Number of full sales	10.63	21.46	0	281
Number of partial sales	1.75	5.98	0	133
Number of Mortgages	2.57	7.48	0	172
Days of atypical temperature	277.24	56.38	96	566
Days of atypical high temperature	157.52	93.52	0	496
Days of atypical low temperature	119.72	90.29	0	564
Number of total allocations	436.52	675.85	0	6,550
Accumulated precipitation	3,539.9	2,836.1	372.2	42,287
	Panel C: Land Registry - Municipality (N = 927)			
Number of owners	2,516.2	2,151.27	18	18,768
Number of plots	2,518.6	2,347.8	17	21,482
Average farm size (ha.)	29.4	94.5	0.65	1,543.5
=1 if land registry update	0.07	0.25	0	1
Registered area (1000 ha.)	39,273.7	84,443.3	170.8	1,465,761
Days of atypical temperature	277.14	56.16	96	566
Days of atypical high temperature	157.68	93.43	0	496
Days of atypical low temperature	119.46	89.67	4	564
Accumulated precipitation	3,488.3	2,804.3	372.2	42,287
	Panel D: ELCA - Household N = 3200			
=1 if HH migrated	0.13	0.33	0	1
=1 if HH has land	0.89	0.31	0	1
=1 if farm size < 3 ha	0.78	0.41	0	1
Farm size (ha.)	2.49	5.54	0	118
Days of atypical high temperature	436.93	165.09	163	816
Days of atypical low temperature	67.03	62.45	0	254
Accumulated precipitation	3792.29	2625.24	720.06	21969.01

Notes: Summary statistics for each estimation sample. Panel A describes the variables used for vereda-level estimations. Total number of sales includes full sales and partial sales during the year. Full sales correspond to sales where the entire property is transferred to another owner. Partial sales correspond to sales that transfer only a fraction of the initial property to a new owner. Number of total allocations corresponds to the cumulative sum of government-allocated plots in the vereda from 1901 until the year of observation. Panel B includes the same information but at municipality level. Panel C summarizes data used for estimations on land distribution at municipality-year level. It takes number of owners, number of plots, average farm size, total registered land and the indicator for land registry update from the national land registry carried out by IGAC. Panel D summarizes data used for estimations at the household-year level. This data comes from 3 rounds (2010, 2013 and 2016) of ELCA, a panel of rural households collected by Universidad de los Andes. Climate data used to compute the number of days with shocks and the accumulated precipitation comes from the Copernicus Climate Change Service (*C3S*). Days with atypical temperature shows the aggregate number of days across the two prior years ($y-2$, $y-1$) with either abnormally high or low temperatures. Accumulated precipitation is the volume of rain in milliliters for year y.

Growing, and South. Within each region, municipalities and veredas were randomly chosen. The baseline sample includes 17 municipalities and 224 veredas. In the follow-up rounds enumerators resurveyed all households and, if the household had split off or migrated, tracked the household head, spouse, and children under nine in 2010. The attrition rate after three waves in 2016 was 13.5%. The household questionnaire collected detailed information on land ownership and migration of household members which we use to complement our empirical analysis. We are interested in how migration, farm size, land ownership, and household consumption change in response to temperature shocks. Panel D of 2.1 contains descriptive statistics of the ELCA panel. On average, 13% of households migrated, 89% had any land and the average size of the plot was 2.5 hectares, 78% of farms are smaller than 3 hectares.

Finally, we study if effects are heterogeneous according to different measures of income and economic conditions of the municipalities. The availability of financial tools like credit access should allow households to smooth consumption without having to sell their property. Similarly, buffer savings and relatively high initial consumption levels (i.e. sufficiently away from a subsistence threshold) should allow households to cope with shocks without having to liquidate their landholdings. Therefore, we expect our results to be stronger in places with higher poverty rates, that are less connected to markets, and that are more isolated and less densely populated. To test this we use municipal-level information collected from CEDE at Universidad de los Andes which consist of a multidimensional poverty measure (the index of Unmet Basic Needs, UBN), a measure of driving distance to the nearest wholesale market, and a rurality index based on measures of population density.[9]

2.3 Empirical strategy

The empirical strategy uses the spatial and temporal variation in the occurrence of adverse weather to estimate the effect of negative productivity shocks on land transactions and the farm size. In

[9] We define municipalities as highly rural if they have a population below 25,000 inhabitants and have a population density below 100 inhabitants per squared kilometer. These thresholds are used by the Colombian government to categorize the 'rurality degree' of municipalities in Colombia. Under this definition close to 63% of municipalities are classified as highly rural.

our first specification we estimate the following equation:

$$s_{v,y} = \beta TempShocks_{v,y} + X'_{v,y}\delta + \eta_v + \kappa_y + \varepsilon_{v,y}, \tag{2.1}$$

where, $s_{v,y}$ is the log number of land sales or mortgages in vereda or municipality v in year y, and $X_{v,y}$ represents a vector of time-varying characteristics composed by rainfall levels in the last three years (y, $y-1$, and $y-2$) and the cumulative number of plots allocated in v from 1901 up to year y. This controls the availability of land for which we can observe transactions.[10] The model includes vereda (or municipality) fixed effects, η_v, that control for time-invariant unobservables, and yearly fixed effects, θ_y, time specific shocks to land markets common to all municipalities. As discussed in section 2.2.3, we define our measure of adverse weather shocks as the sum of days with atypical temperatures (denoted as $AtypicalDay_{v,d}$) in the two years prior:

$$TempShocks_{v,y} = \sum_{s=y-2}^{y-1} AtypicalDay_{v,s}. \tag{2.2}$$

Both the model in equation (2.1) and all subsequent specifications rely on the identifying assumption that there are no vereda- or municipality-specific, time-varying unobservable characteristics correlated to the occurrence of atypical weather events, i.e., conditional on the set of fixed effects the occurrence of temperature shocks is as good as random; a standard assumption in the literature (see e.g., Dell et al. (2014)). In section 2.4.2, we show that our result are robust to specifications that include in addition state-specific time trends and to alternative measures of atypical temperature computed using fixed thresholds. We cluster standard errors in all regressions at the municipality level.

To measure the effect on the distribution of farm sizes we first estimate a model analogous to the one in equation (2.1) above but using the land registry data. We estimate for municipality m

[10] Regressions where the dependent variable is instead defined as the number of sales divided by cumulative allocations yields qualitatively identical results.

and year y, the model:

$$n_{m,y} = \rho TempShocks_{m,y} + X'_{m,y}\nu + \mu_m + \kappa_y + \epsilon_{v,y}, \tag{2.3}$$

where, $n_{m,y}$ is either the log number of land plots or land owners, or the log average or median areas of plots and areas per owner in municipality m in year y.[11] The vector of controls $X_{m,y}$ contains rainfall levels in the past three years, a dummy indicating if there was a cadastral update in the municipality that year, and the log of total municipal land area recorded in the registry. Municipality and year fixed effects are represented by μ_m and κ_y, respectively.

While the model in equation (2.3) allows us to estimate how productivity shocks have an effect on different moments of the municipal farm-size distribution, it is not informative on whether these changes are driven by the sale and transfer of farms of a specific size. For example, a reduction in the average farm size within a municipality could be equally driven by the fragmentation of large estates into medium-sized farms without there being any change in the number of small farms, as by the fragmentation of small farms into even smaller ones without having any change in the number of larger properties.

In order to investigate the type of farm size where the effect of negative productivity shocks translates more strongly into property transfers, we estimate how the number of owners within fixed farm-size bins changes across time. We do this by splitting the distribution of farm sizes within each municipality by quantiles, such that each quantile has, in the initial year of our sample, the same number of farm owners.[12] Keeping these quantile thresholds fixed, we then compute for each subsequent year the number of owners within each bin. If, for example, average farm sizes are dropping due to the partition of the largest plots, we would then observe a sharp reduction in the number of owners with landholding areas at the –fixed– top quantile of the initial farm-size distribution.

Denote as $\{q_m^1, ..., q_m^J\}$ the areas defining each of the j quantiles of farm size distribution in

[11] We define *plots* as a piece of land with a distinct registry number, an owner can have several –not necessarily contiguous– plots.

[12] We take the initial distribution to be the year 2000, for which 97% of municipalities have registry information. For the remaining municipalities, we take the initial distribution to be the one observed in the first year in which they appear in the land registry dataset.

municipality m in the year 2000, and denote as $AreaOwned_{i,m,y}$ the total landholdings of farmer i in municipality m on year y. We compute for each year the number of owners with total landholdings within each of these fixed size bins as:

$$NumOwners_{q_m^j,m,y} \equiv \sum_{i \in m} \mathbb{1} \cdot [AreaOwned_{i,m,y} \in (q_m^{j-1}, q_m^j)], \tag{2.4}$$

where $j = 1, ..., J$, and $q_m^0 = 0$ for all m. We use this variable to estimate independent regressions (one per quantile j) of the form:

$$NumOwners_{q_m^j,m,y} = \gamma^j TempShocks_{m,y} + X'_{m,y}\xi^j + \mu_m^j + \kappa_y^j + \omega_{v,y}^j, \tag{2.5}$$

where all the right-hand-side variables are the same as in

We finally estimate household-level regressions, using data from the ELCA survey, to investigate the effect of adverse weather shocks on household's decisions. We estimate the model:

$$h_{i,v,y} = \alpha TempShocks_{v,y} + X'_{v,y}\tau + \iota_i + \kappa_y + \psi_{v,y}, \tag{2.6}$$

where $y = \{2010, 2013, 2016\}$, and $h_{i,m,y}$ is either log per capita consumption, a dummy indicating household migration, different measures of land ownership, or measures of work outside agriculture. $X_{v,y}$ represents rainfall levels in the past three years, ι_i represents household-level fixed effects and κ_y year fixed effects.

2.4 Results

2.4.1 Reduced-form results

Table 2.2 presents the OLS estimates from equation (2.1) on our four measures of land transactions. Columns 1 and 5 report the effect of weather shocks on all types of land sales within veredas and municipalities respectively, while columns 2 and 6 report the effect on sales that transfer the entire area of a plot to the new owner, which we denote as 'full' sales. Columns 3 and 7 report the effect on partial sales. Increases in the frequency of adverse weather shocks raise the number of land transactions. This result holds regardless of whether the observation unit is set at the municipality or at the vereda level. Land sales caused by adverse shocks are entirely driven by full sales when the unit of observation is set at the vereda level. By contrast, when observed at the municipality level, the effect on partial sales is substantially higher. This disparity in the effect of shocks on partial land sales might be related to unobserved characteristics related to the selected nature of the vereda sample. For example, veredas with better record keeping practices which we are thus better able to match in the data might be also richer or situated closer to urban centers. These characteristics could also be the reason why the effect of shocks on partial sales (and more generally on all types of transactions) is smaller than when compared to the –unselected– municipality sample. Consistent with this hypothesis, results shown in section 2.4.3 below do indicate that shocks have a stronger effect on transaction frequency in less densely populated and more isolated regions.

Table 2.2: Temperature Shocks and Land Sales

	Vereda level panel				Municipality level panel			
	Total (1)	Full (2)	Partial (3)	Mortg. (4)	Total (5)	Full (6)	Partial (7)	Mortg. (8)
$TempShocks_{v,y}$	0.020***	0.022***	0.003	0.022***	0.076***	0.088***	0.116***	0.104***
	(0.006)	(0.006)	(0.005)	(0.006)	(0.021)	(0.023)	(0.028)	(0.020)
Observations	149,664	149,664	149,664	149,664	10,392	10,392	10,392	10,392
R-Squared	0.574	0.561	0.359	0.392	0.912	0.903	0.710	0.793
Mean Dep. Var.	0.55	0.47	0.07	0.11	12.38	10.63	1.75	2.57

Notes: Data from the National Superintendency of Notaries (SNR) records. Columns 1 and 5 show the effect on total (full + partial land sales) columns 2 and 6 show the effect on full sales (when the entire property is transferred to another owner), columns 3 and 7 show the effect on partial sales (when only a fraction of the plot is transferred), and columns 4 and 8 show the effect on mortgage originations. All dependent variables are in log(x+1) transformation. The main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions also include year and geographic fixed effects (vereda or municipality). *Mean Dep. Var.* is the mean of the untransformed variable. Standard errors clustered at the municipality level reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Columns 4 and 8 of Table 2.2 show that adverse shocks lead to a substantial increase in the number of mortgages taken out by farmers against their properties. In the case of the municipality sample, the magnitude of the effect on mortgages is roughly 30% larger than on total sales. While the use of mortgages is uncommon in rural Colombia (our data shows that on average only 2.6 mortgage originations happen in a municipality per year) this result clearly indicates that weather shocks lead farmers to look for ex-post mechanisms that allow them to cope, and that it is in the absence of such mechanisms that land sales might become a last resort measure. Indeed, heterogeneity results shown in section 2.4.3 show that in richer municipalities mortgages as a response to shocks are roughly twice more likely to occur than sales, while the opposite is true in poorer, more isolated municipalities.

The increases in the frequency of land sales caused by weather shocks further translates into a reduction in average farm sizes. Table 2.3 presents the results of estimating equation (2.3) on different measures of municipal land size using the land registry data. More days of atypical temperature in a municipality during the previous two years lead to an increases the number of plots and owners (columns 1 and 2), and thus to lower average farm and plot sizes (columns 3 and 4). Taken together, the magnitudes of these effects are economically important and suggest that the presence of uninsured covariate shocks play an important role in determining land distribution patterns. An additional 100 days of atypical temperature (roughly a two standard deviation increase) throughout a two-year period increase the number of land purchases and mortgage originations in a municipality by 7.6% and 10.4% respectively, while reducing the average farm size by 1.2%.

The results shown in Table 2.3 suggest that the net effect of weather shocks on land distribution patterns is to increase fragmentation. However it is not possible to know from that estimation alone if there is a specific part of the farm size distribution responsible for the overall decrease in average area owned. In order to investigate this, we estimate equation (2.5) on 10 quantiles of the initial municipality-level farm size distribution. The coefficients of interest from these regressions are summarized in Figure 2.4. Negative weather shocks cause a sizable increase in the number of owners with farms on the lower 5 deciles of the initial distribution, but no statistically significant effect on the number of owners in the 5 top deciles.[13] This result shows that the observed reduction in mean farm sizes caused by weather shocks is entirely driven by the subdivision and sale of smaller farms to new owners that did not have any additional landholdings. The fact that there is no noticeable change in the number of owners in the right part of the initial distribution indicates

[13] Regression results in table form are in 2.7 in the appendix. Appendix Figure 2.10 shows analogous estimations for alternative partitions ($j = 5$, and $j = 20$) of the initial farm size distribution.

that large landholders are not driven to sell their land after facing a weather shock. This result is not surprising under the presumption that large landholders are more likely to have buffer savings and better access to credit than small farmers. However, these results also show that large landholders fail to use the expansion in land supply caused by adverse weather shocks to increase their own landholdings (a fact shown in Figure 2.4 but, more generally, evidenced as well in the previous set of results which show that average landholding area falls).

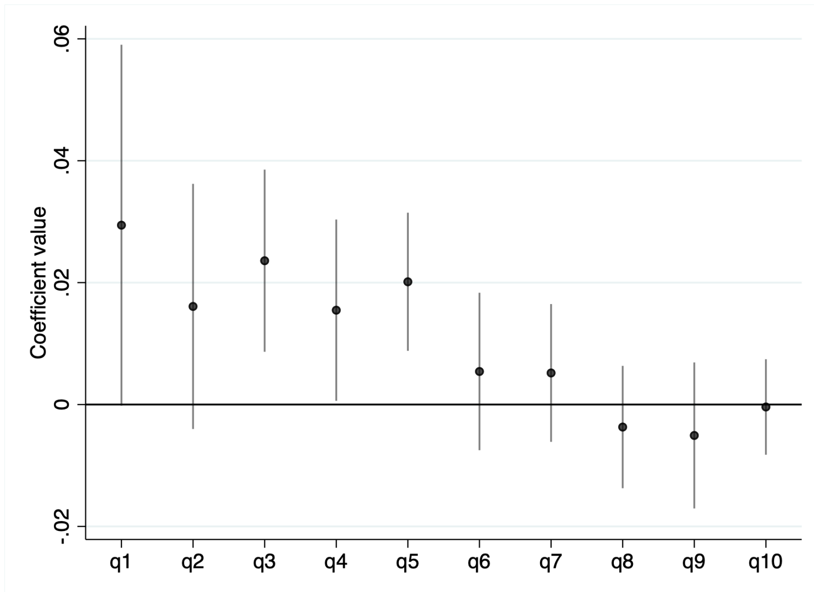
Table 2.3: Temperature Shocks and Average Farm Size

	Number of Plots	Number of Owners	Mean Plot Size	Mean Area/Owner	Median Plot Size	Median Area/Owner
	(1)	(2)	(3)	(4)	(5)	(6)
$TempShocks_{v,y}$	0.0120**	0.0120***	-0.0120**	-0.0123***	-0.0164	-0.0126
	(0.0048)	(0.0045)	(0.0048)	(0.0046)	(0.0113)	(0.0089)
Observations	10,934	10,934	10,934	10,934	10,934	10,934
R-squared	0.9905	0.9920	0.9935	0.9947	0.9763	0.9881
mean.dep.var	2519	2516	30.50	29.36	15.22	12.88

Notes: Data from the National Land Registry (*Catastro Nacional*), mantained by the National Geographical Institute (IGAC). All dependent variables are in logarithms. The main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions also include year and geographic fixed effects (vereda or municipality). *Mean Dep. Var.* is the mean of the untransformed variable. Standard errors clustered at the municipality level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

We also use data from the ELCA survey to explore if observed household-level decisions in response to shocks are consistent with the aggregate patterns on land sales and farm size distribution we document. Table 2.4 presents the results of estimating equation (2.6) on several household-level variables for years 2010, 2013, and 2016. Column 1 shows that more days of atypical temperature increase the probability that the household migrates, a result that is consistent with Ibáñez et al. (2022), who show that households migrate in el Salvador in response to temperature shocks. Column 2 shows an imprecisely estimated negative effect of shocks on the size of the household farms, but columns 3 and 4 do show evidence that shocks lead households to liquidate their landholdings and increase the likelihood that the household farm has less that 3 hectares of land. Column 5 shows that a 100 day increase in the number of days with harmful temperatures increases the probability of a household head shifting from agricultural to non-agricultural activities by 7.7%, while column 6 shows that there are no statistically significant effects on the probability that the household head works off farm. Finally, column 7 shows that weather shocks have a sizable effect on the monetary value of per-capita consumption –a 12.2% drop per 100 additional days–. This result suggest that households are not able to fully smooth consumption.

Figure 2.4: Temperature Shocks and Number of Owners by Initial Distribution Quantiles



Notes: OLS estimates of the γ coefficients according to equation (2.5), for each of the 10 quantiles of the initial municipality-level distribution of farm sizes. Each point estimate corresponds to a separate regression where the main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. The dependent variable is number of owners per quantile and is in logarithms. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions include year and municipality fixed effects. Error bars display 95% confidence intervals for standard errors clustered at the municipality level.

These micro-level responses to weather shocks are broadly consistent with the distribution-wide effects on farm sizes presented above. Note that ELCA was designed to cover and be representative of small agricultural producers, so the fact that we find that these households are migrating and reallocating labor away from agriculture falls in line with the result from Figure 2.4 showing that it is mostly farmers on the lower tail of the farm size distribution the ones who respond to adverse shocks by asset liquidation.

2.4.2 Robustness

Appendix tables 2.8 and 2.9 present the results of two robustness checks that we perform. First, we estimate our baseline specifications in equations 2.1 and 2.3 including state-specific time trends. This allows us to rule out that spurious correlations between regional time trends of temperature

Table 2.4: Temperature Shocks and Household Decisions

	Household Migrated	Farm Size	Household Has Land	Farm Size $\leq$ 3 ha	Sector Not Agri.	Work off Farm	Consumption per capita
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$TempShock_{v,y}$	0.064***	-0.126	-0.050***	0.049***	0.077**	-0.010	-0.122***
	(0.019)	(0.088)	(0.016)	(0.019)	(0.034)	(0.023)	(0.026)
Observations	12,124	10,756	11,987	12,124	7,523	12,124	10,884
R-squared	0.555	0.779	0.678	0.717	0.767	0.537	0.729
Mean Dep. Var.	0.107	2.875	0.900	0.777	0.242	0.749	2.665

Notes: Data from ELCA. Dependent variables are, from left to right: a dummy indicating if household migrated between survey waves; area owned by the household in hectares; a dummy indicating if household owns any land; a dummy indicating if household's landholdings are below 3 hectares; a dummy indicating if household head works in the non-agricultural sector; a dummy indicating if household head main economic activity happens outside the family farm; value of per capita consumption in 2016 colombian pesos (in millions). All regressions include a control of aggregate rainfall and household-level fixed effects. *Mean Dep. Var.* is the mean of the untransformed variable. Robust standard errors reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

shocks and our variables of interest are driving our results. As shown in Panel A of both tables, the coefficient estimate on land sales (Table 2.8) and farm size (Table 2.9) remain unchanged.

Second, we estimate equations 2.1 and 2.3 using an alternative measure of temperature shocks. In particular, we follow Aguilar-Gomez et al. (2022), and we compute countrywide temperature thresholds for atypically high and low temperatures. We use the distribution of maximum and minimum daily temperatures across all municipalities over the sample period and define high and low-temperature thresholds as the 95^{th} and the 5^{th} percentiles of maximum and minimum temperatures, respectively. As in our main specifications, we then add the total number of days with temperature above (below) the high (low) threshold over a two-year window. Panel B of tables 2.8 and 2.9 show that our main results on land sales and farm size are robust to this alternative definition of temperature shock.

2.4.3 Heterogeneity

In this section we explore the potential mechanisms that might drive our results on land sales and farm size. Table 2.2 shows that weather shocks have a sizable impact on the number of landholders that access credit by using their land as collateral. Being able to mortgage property is not likely, however, to be accessible for most landholders in poorer and more isolated economies where credit markets are less developed. Incomplete credit markets could therefore be a potential driver behind our findings. Similarly, if households need to meet a minimum subsistence consumption threshold, the ability to cope with drops in income by cutting back on expenses is more reduced the closer the initial consumption levels are to the subsistence threshold. Faced with a shock, poorer households should be then more likely to be forced to liquidate their assets in order to maintain a minimum consumption level. We would expect households in poorer municipalities to respond more strongly to shocks through land sales and, by contrast, we would expect households in richer municipalities to respond more strongly through mortgage originations.

We show in Table 2.5 the result of estimating equation (2.1) with an additional interaction term indicating if a municipality is i) above the median in a multidimensional poverty index calculated by the national government, ii) above the median in the distance required to reach a wholesale market, and iii) above the median in a 'rurality' index measuring low population density. Consistent with the hypothesis of credit constraints, results in column 4 show that the positive effect of shocks on mortgage originations in high poverty municipalities is roughly three times smaller than the size of the effect in low poverty municipalities. Similarly, columns 8 and 12 show that this same effect on mortgages in municipalities with above-median distance to wholesale markets or with low population density is roughly half the size of the effect observed in municipalities with stronger market access and higher population densities. We take these results as suggestive evidence of the potential for credit markets to prevent distress sales.

Table 2.5: Temperature Shocks and Land Sales - Heterogeneous Effects

	H_i: High Multipoverty Index				H_i: High Distance to Market				H_i: Low Population Density			
	(1) Total	(2) Full	(3) Partial	(4) Mortgage	(5) Total	(6) Full	(7) Partial	(8) Mortgage	(9) Total	(10) Full	(11) Partial	(12) Mortgage
$TempShocks_{v,y}$	0.0894***	0.102***	0.142***	0.169***	0.0727***	0.105***	0.0872***	0.151***	0.0403	0.0574**	0.0946***	0.161***
	(3.66)	(3.96)	(4.47)	(6.82)	(2.96)	(3.99)	(2.84)	(6.29)	(1.56)	(2.04)	(2.86)	(6.52)
$TempShocks_{v,y} \times H_i$	-0.0184	-0.0226	-0.0275	-0.113***	0.00560	-0.0291	0.0502*	-0.0828***	0.0554**	0.0477*	0.0330	-0.0898***
	(-0.77)	(-0.95)	(-0.96)	(-5.01)	(0.24)	(-1.24)	(1.84)	(-3.77)	(2.28)	(1.92)	(1.16)	(-4.03)
Observations	9924	9924	9924	9924	10392	10392	10392	10392	10392	10392	10392	10392
R-Squared	0.913	0.904	0.710	0.794	0.912	0.903	0.711	0.794	0.912	0.903	0.710	0.794

Notes: Data from the National Superintendency of Notaries (SNR) records. Columns 1 and 5 show the effect on total (full + partial land sales) columns 2 and 6 show the effect on full sales (when the entire property is transferred to another owner), columns 3 and 7 show the effect on partial sales (when only a fraction of the plot is transferred), and columns 4 and 8 show the effect on mortgage originations. All dependent variables are in log(x+1) transformation. The main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions also include year and geographic fixed effects (vereda or municipality). Standard errors clustered at the municipality level reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

The previous results shed light on poverty and credit constraints as potential mechanisms driving distress sales. These, however, do not necessarily imply land fragmentation. For example, it is possible that large landowners, who are likely to be less credit-constrained, buy land from poor farmers after the shock. This, contrary to our findings, could lead to land consolidation. We now explore if transaction costs that result from the lack of contiguity between large and small farms hinder this process of land consolidation and drive the decreases in farm size that we observe in the data.

We construct two measures of contiguity between large and small farms in a municipality. First, we use land registry maps, available for 2017, to compute the share of plots below the 10^{th} percentile of the size distribution in the municipality that are contiguous to at least one plot above the 90^{th} percentile of this distribution. We then classify municipalities with high contiguity as those with a share above the national median. We also carry out the same exercise using as alternative thresholds the 20^{th} and 80^{th} percentiles of the distribution.

Our second measure uses the GPS coordinates of the farms included in the 2013 agricultural census. We compute buffers around this GPS coordinates to simulate 1.2 times the area of each farm and define as contiguous two farms with overlapping buffers. Using this measure, we compute the share of small farms that are contiguous to at least one large farm. As before, we define small farms as those below the 10^{th} percentile of the municipality size distribution and large farms as those above the 90^{th} percentile, and we conduct the same analysis varying these thresholds to the 20^{th} and 80^{th} percentiles respectively.

Table 2.6 presents the results of estimating equation 2.3, with the interaction between the temperature shocks and farm contiguity. Panel A presents the results with the measure we compute with the registry maps. Panel B uses the measure of the overlapping buffers. Odd-numbered columns show results for the 90^{th}-10^{th} percentile definition of large and small plots, while even-numbered columns show results for the 80^{th}-20^{th} percentile definition. As shown in both panels, the coefficient estimates for the temperature shocks have the same sign and statistical significance as in our baseline specification. The interaction term is small in magnitude and not statistically significant regardless of the measure of contiguity. This result suggests that transaction costs associated with the lack of contiguity between large and small farms might not be driving land fragmentation in this context.

Table 2.6: Temperature Shocks and Farm Size - Heterogeneous Effects

	Number of Plots		Numbers of Owners		Mean Plot Size		Mean Area/Owner	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Land Registry Map - Contiguous Plots							
$TempShocks_{v,y}$	0.011**	0.011*	0.010*	0.009*	-0.011**	-0.011*	-0.010**	-0.009*
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
$TempShocks_{v,y} \times High$	-0.005	-0.004	-0.004	-0.002	0.005	0.004	0.004	0.001
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
Observations	10,413	10,448	10,413	10,448	10,413	10,448	10,413	10,448
R-squared	0.990	0.990	0.992	0.992	0.994	0.994	0.995	0.995
Mean Dep. Var	2,576.47	2,568.60	2,582.51	2,574.54	30.41	30.40	29.16	29.16
	Panel B: Agricultural Census Coordinates - Overlapping Buffers							
$TempShocks_{v,y}$	0.016***	0.018***	0.016***	0.017***	-0.016***	-0.018***	-0.017***	-0.017***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)
$TempShocks_{v,y} \times High$	-0.006	-0.009	-0.003	-0.004	0.006	0.009	0.003	0.004
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
Observations	9,402	9,402	9,402	9,402	9,402	9,402	9,402	9,402
R-squared	0.990	0.990	0.992	0.992	0.993	0.993	0.995	0.995
Mean Dep. Var	2,552.60	2,552.60	2,548.64	2,548.64	29.27	29.27	28.23	28.23

Notes: Data from the National Land Registry (*Catastro Nacional*), mantained by the National Geographical Institute (IGAC). All dependent variables are in logarithms. The main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions also include year and geographic fixed effects (vereda or municipality). Panels A and B use land registry maps and agricultural census data, repsectively, to measure plot contiguity. Odd-numbered columns in both panels show results for the 90^{th}-10^{th} percentile definition of large and small plots, while even-numbered columns show results for the 80^{th}-20^{th} percentile definition. See text for more details. *Mean Dep. Var.* is the mean of the untransformed variable. Standard errors clustered at the municipality level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

2.5 Model

To rationalize the reduced-form results in Section 3.4, we now develop a model of agricultural production in which heterogeneous agents own and transact land. A key feature of the model is that land sales serves as a consumption smoothing device, which is triggered by subsistence consumption motives when the agricultural sector is subject to a negative weather shock. We next present the environment of the model, we then simulate the impact of negative weather shocks in the economy.

2.5.1 Environment

Consider a small-open economy region with two sectors, agriculture and non-agriculture, that operates over two periods t_1, t_2. There are a measure of agents N who are heterogeneous in terms of their endowments of two distinct assets: land (l_0) and wealth (m_0). In each period, agents choose whether to be a worker in the non-agricultural sector or a farmer.

Output in the agricultural sector employs land:

$$y_t(l) = a_t l$$

where $y_t(l)$ is the total output, a_t is the productivity of land and l is the employed land. The productivity parameter a_t is subject to two weather conditions, high a_H and low L ($a_L > a_H$). If agents choose to be a farmer, they earn the agricultural output of their landholdings $y_t(l)$. If they choose to become a worker, they earn wages $w_t = w$. Moreover, in each period, agents earn the returns from their wealth $r_t m_t(\omega)$, where we assume an exogenous return $r_t = r$. Preferences are Stone-Geary:

$$U = \sum_{t=\{1,2\}} \log\left(c_t - c_S\right)$$

where c_S is a subsistence consumption level. Budget constraints for each period are:

$$f_t\left[a_t l_t\right] + (1 - f_t)\left[w_t\right] + r_t m_{t-1} = c_t + m_t + p_t\left(l_t - l_{t-1}\right)$$

where f_t is an indicator function equal to 1 if the agent choose to be a farmer and 0 otherwise. On the left hand side we have the earnings of the agent in period t. On the right hand side we have the expenditure of agents. Agents choose how much to consume (c_t), how much wealth to acquire (m_t), and how much land to own (l_t). If they choose to own more land than the one inherited from $t-1$—i.e., if $l_t - l_{t-1} > 0$—, they must purchase land by a price p_t. Conversely, they can sell land and collect $p_t\left(l_t - l_{t-1}\right)$—if $l_t - l_{t-1} < 0$. We impose two borrowing constraints:

$$0 \leq l_t$$

$$0 \leq m_t.$$

In other words, agents are unable to sell more land than they inherited. In addition, agents are unable to borrow wealth.

The timing of the model is as follows. First, agents observe the weather conditions for the next two periods, their initial endowments of land l_0 and wealth m_0, and choose whether to become a worker or a farmer. (For simplicity, we assume that agents can choose only one occupation for the two period, so that $f_1 = f_2$.) Then, in each period t, they collect their earnings from work (w_t), from land ($a_t l_t$), and from accumulated wealth ($r_t m_{t-1}$), and choose how much to consume.

2.5.2 Equilibrium

We define the market equilibrium as follows. Given a distribution of land and wealth endowments $\{l_0, m_0\}$, a total land a sequence of weather events $\{a_t\}$, wages $\{w_t\}$, and returns to wealth $\{r_t\}$, the market equilibrium is a sequence of consumption, wealth, land ownership and land prices, $\{c_t^*, l_t^*, m_t^*, p_t^*\}$ such that all agents make optimal choices and the following market clearing condition holds

$$\int l_t(\omega)dF(\omega) = \int l_{t-1}(\omega)dF(\omega)$$

where ω indexes an agent and $F(\omega)$ is the distribution of agents.

2.5.3 Land transactions under adverse weather shocks

To observe any land transaction between agents, it must be the case that the wealth asset has an intermediate rate of return that lies in-between the low agricultural productivity a_L in $t = 1$ and the high productivity a_H in $t = 2$. We thus impose the conditions: $r_1 \geq a_L$, and $r_2 < a_H/(p_1 - a_L)$. Under these conditions agents who decide to become farmers will hold all of their wealth in land, and conversely agents who decide to become non-farm workers will hold all of their wealth in the alternative asset.[14]

Optimal input demands for farmers are then:

$$\begin{aligned} l^*_{1,F} &= \frac{1}{2(p_1 - a_L)}\left[(2a_L - p_1)l_0 + r_1 m_0 + \frac{(p_1 - a_H - a_L)}{a_H}c_S\right] \\ m^*_{1,F} &= 0, \end{aligned}$$

which yields utility

$$U^*_F = \log\left(a_L l_0 + r_1 m_0 - (p_1 - a_L)\, l^*_{1,F} - c_S\right) + \log\left(a_H\left(l_0 + l^*_{1,F}\right) - c_S\right).$$

For their part, optimal input demands for workers are:

$$\begin{aligned} l^*_{1,W} &= -l_0 \\ m^*_{1,W} &= \frac{1}{2r_2}\left[r_1 r_2 m_0 + r_2 p_1 l_0 + (1 - r_2)(c_S - w)\right] \end{aligned}$$

[14]The knife-edge case where $r_2 = a_H/(p_1 - a_L)$ does allow for agents simultaneously demanding positive amounts of both assets. For ease of exposition we abstract away from this case.

with utility

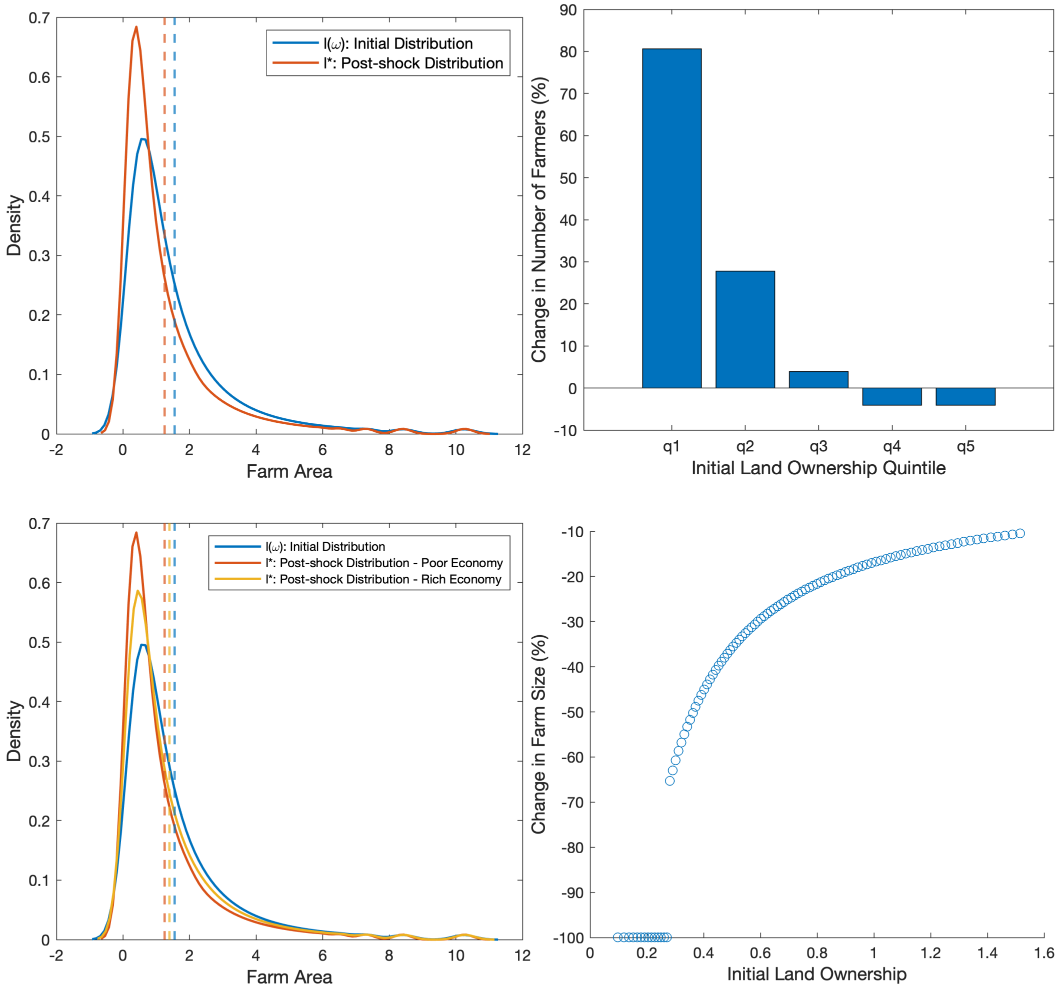
$$U_W^* = \log\left(r_1 m_0 + w - m_{1,W}^* - c_S\right) + \log\left(w + r_2 m_{1,W}^* - c_S\right).$$

An agent chooses to become a farmer if $U_F^* \geq U_W^*$.

2.5.4 Simulation results

We simulate the economy described above and document the changes in the equilibrium farm size distribution with respect to initial endowments. Results from this exercise are shown in Figure 2.5. The simulated economy is able to replicate the main reduced-form results presented in Section 3.4. First, as shown in the upper-left panel of Figure 2.5, the initial shock to productivity in $t = 1$ induces a net increase in the number of agents occupied in the agricultural sector and a decrease in the average size of individual farms. Second, also consistent with our empirical findings, the upper-right panel of the same figure shows that this result is almost entirely driven by the increase in the number of agents operating smaller-sized farms, with almost no variation in the number of landholders operating farms in the upper quintiles of the initial size distribution. Third, the bottom-right panel of the figure also shows how, despite an observed net-increase in the number of farmers, the productivity shock does lead a fraction of the smallest initial landholders to exit agriculture altogether. Finally, the bottom-left panel shows that the observed effects of the shock on the farm size distribution are more pronounced in underdeveloped economies, where we parameterize an economy to be 'poorer' if it has a relatively higher subsistence constraint.

Figure 2.5: Change in Farm Size Distribution - Simulation Results



Notes: Simulated distribution of equilibrium farm sizes with respect to initial endowments. The simulated economy consists of 300 agents with asset endowments drawn from a log-normal distribution with mean = 0 and standard deviation = 1. Top Left Panel: Blue solid line shows the distribution of land endowments before trade occurs. Red solid line shows the equilibrium farm size distribution in $t = 2$. Dashed lines represent average farm sizes. Top Right panel: Percentage change in the number of agents occupied in agriculture by quintiles of the initial land distribution. Bottom Left Panel: Blue solid line shows the distribution of land endowments before trade occurs. Red solid line shows the equilibrium farm size distribution in $t = 2$ for a ‘’poor’ economy parameterized as having a high subsistence constraint $c_S = 0.45$. Yellow solid line shows the equilibrium farm size distribution in $t = 2$ for a ‘rich’ economy parameterized as having a low subsistence constraint $c_S = 0.15$. Dashed lines represent average farm sizes. Bottom Right Panel: Percent change in landholdings after trade for agents initially endowed with land. Additional parameter values: $a_L = 1$; $a_H = 5$; $r_1 = 1$; $r_2 = 1$; $c_S = \{0.45, 0.15\}$; $w = 0.5$

2.5.5 Discussion

A key result of the model is that a fraction of agents initially endowed with land will exit agriculture despite the absence of uncertainty, and the fact that all agents are aware that productivity will be higher in the second period. This behavior is driven instead by the fact that agents with the lowest land endowments run against the subsistence consumption constraint and are forced to forego higher consumption in the future to achieve minimum consumption levels in the present. This need to smooth consumption across periods leads to an increase in aggregate land supply and an initial depression of land prices. Faced with the opportunity to acquire relatively cheap assets, both large farmers and agents not initially endowed with land have incentives to increase their landholdings. Under similar conditions, any version of a farm-production model that lacks a non-farm sector will necessarily yield as a result the consolidation of small farms into larger holdings and an increase in average farm size. By contrast, the model presented above is capable of yielding land fragmentation as a result due to the fact that the non-farm sector is unaffected by the initial slump in productivity. Being isolated from the shock, and thus becoming richer relative to landholders, agents endowed with large amounts of the alternative 'wealth' asset are able to outbid large landholders for the excess land supply and enter the soon-to-be more profitable agricultural sector.

An important feature of this version of the model is that agent heterogeneity is solely driven by differences in initial endowments and not by differences in relative productivity. This entails that, as it stands, all observed changes in asset ownership and on the operational scale of farms have no implications for the aggregate productivity of the economy. A more comprehensive version of this model that allows for heterogeneous productivity across farmers might be able to produce richer predictions related to specific selection effects caused by shocks and the consequences these can have on the aggregate efficiency of the agricultural sector.

More generally, and in order to gauge both the impact of the increasing risk of extreme weather shocks induced by climate change on agricultural productivity, as well as the aggregate impact of the expansion of rural-credit or agricultural insurance programs, we further plan to develop and structurally estimate a model of agricultural production that combines the farm production structure developed in Gáfaro and Pellegrina (2022) with the borrowing frictions modeled in the macro-literature on heterogeneous agent models (Krusell and Smith, 1998; Buera et al., 2011). Our goal will be to inform the behavioral parameters of the model driving agents' decisions to sell and buy agricultural land, and derive a set of results that can indicate under which conditions the model rationalizes the qualitative features of the data. These behavioral parameters should then allow us to estimate policy-relevant counterfactuals of interest that shed light on the potential effects that future increases in weather shock frequency and severity will have on land-distribution patterns and agricultural productivity in developing economies.

2.6 Conclusion

This paper explores the effect of uninsured weather shocks on distress sales and the farm size in Colombia. Exploiting a unique combination of datasets that include the transaction history of hundreds of thousands of individual plots and a municipal-level census of rural properties we find that shocks lead to an increase in the frequency of land sales and to a reduction in average farm size. This reduction is driven by the smaller farms in the initial farm-size distribution being further subdivided and purchased by previously landless individuals. Consistent with the aggregate patterns we find on land sales and land distribution, we also show that these shocks decrease household consumption and induce rural households to migrate, engage in non-agricultural activities, and operate smaller farms.

Distress sales after a negative covariate productivity shocks might depress land prices. However, a standard heterogeneous-agent model with credit market imperfections would predict that this excess supply of land should lead to the consolidation of many small farms into larger landholdings. Our results are at first glance puzzling since we show that the opposite effect, land *fragmentation*, takes place. We rationalize the results with a model where agents have to make an intertemporal consumption decision while facing a minimum subsistence constraint, and heterogeneity in initial endowments causes some agents to be isolated from the initial negative productivity shock. The combination of the shock and the subsistence constraint induces an expansion of the aggregate supply of land; the presence of relatively wealthier landless agents who find themselves unaffected by the shock and who can profit from the temporary drop in land values then leads to a net increase in the number of agents occupied in agriculture and to lower average farm sizes.

Our empirical findings could be explained by an alternative model where, for example, frictions on land assembly stemming from the potential non-contiguous character of land plots for sale are present (e.g. Brooks and Lutz (2016)). Identifying the specific mechanisms that prevent land from becoming endogenously consolidated would greatly improve our understanding on the organization of economic activity in the agricultural sector of much of the developing world. The evidence that we present in this paper suggests that uninsured weather shocks constitute a substantial barrier for productivity improvements in the agricultural sector of developing countries. Given that extreme temperature shocks are expected to increase in frequency and severity in the near future, these findings have important policy implications related to the expansion of financial tools designed for risk management in rural settings.

Bibliography

Adamopoulos, T. and Restuccia, D. (2014). The Size Distribution of Farms and International Productivity Differences. *American Economic Review*, 104(6):1667–1697.

Adamopoulos, T. and Restuccia, D. (2020). Land reform and productivity: A quantitative analysis with micro data. *American Economic Journal: Macroeconomics*, 12(3):1–39.

Aguilar-Gomez, S., Gutierrez, E., Heres, D., Jaume, D., and Tobal, M. (2022). Thermal stress and financial distress: Extreme temperatures and firms' loan defaults in mexico. *Available at SSRN 3934688*.

Albertus, M. (2015). *Autocracy and Redistribution: The Politics of Land Reform*. Cambridge University Press, Cambridge.

Aragón, F. M., Oteiza, F., and Rud, J. P. (2021). Climate change and agriculture: subsistence farmers' response to extreme heat. *American Economic Journal: Economic Policy*, 13(1):1–35.

Arteaga, J., Osorio, C. C., Cuéllar, D., Ibáñez, A. M., Botero, L., Murcia, M., Neva, J., and Nieto, Á. (2017). Fondo de Tierras del Acuerdo Agrario de La Habana: Estimaciones y propuestas alternativas. *Documento CEDE*, No. 41:108.

Brooks, L. and Lutz, B. (2016). From today's city to tomorrow's city: An empirical investigation of urban land assembly. *American Economic Journal: Economic Policy*, 8(3):69–105.

Buera, F. J., Kaboski, J. P., and Shin, Y. (2011). Finance and development: A tale of two sectors. *American economic review*, 101(5):1964–2002.

Burke, M., Hsiang, S. M., and Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, 527(7577):235–239.

Cain, M. (1981). Risk and insurance: Perspectives on fertility and agrarian change in india and bangladesh. *Population and development review*, pages 435–474.

Carter, M., de Janvry, A., Sadoulet, E., and Sarris, A. (2017). Index insurance for developing country agriculture: a reassessment. *Annual Review of Resource Economics*, 9:421–438.

Carter, M. R. and Zimmerman, F. J. (2003). Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints. *Journal of Development Economics*, 71(2):233–260.

Caselli, F. (2005). Accounting for cross-country income differences. *Handbook of economic growth*, 1:679–741.

Chen, C., Restuccia, D., and Santaeulàlia-Llopis, R. (2022). The effects of land markets on resource allocation and agricultural productivity. *Review of Economic Dynamics*, 45:41–54.

CNMH (2016). *Tierras y conflictos rurales: historia, políticas agrarias y protagonistas*. Informes de investigación. Centro Nacional de Memoria Histórica.

Cole, S., Giné, X., and Vickery, J. (2017). How does risk management influence production decisions? evidence from a field experiment. *The Review of Financial Studies*, 30(6):1935–1970.

Colmer, J. (2021). Rainfall variability, child labor, and human capital accumulation in rural ethiopia. *American Journal of Agricultural Economics*, 103(3):858–877.

Deininger, K. and Jin, S. (2008). Land sales and rental markets in transition: Evidence from rural vietnam. *Oxford bulletin of Economics and Statistics*, 70(1):67–101.

Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? the new climate economy literature. *Journal of Economic Literature*, 52(3):740–798.

Fafchamps, M. (1992). Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World. *American Journal of Agricultural Economics*, 74(1):90–99.

Faguet, J.-P., Sánchez, F., and Villaveces, M.-J. (2020). The perversion of public land distribution by landed elites: Power, inequality and development in Colombia. *World Development*, 136.

Foster, A. D. and Rosenzweig, M. R. (2022). Are there too many farms in the world? labor market transaction costs, machine capacities, and optimal farm size. *Journal of Political Economy*, 130(3):636–680.

Gáfaro, M. and Pellegrina, H. S. (2022). Trade, farmers' heterogeneity, and agricultural productivity: Evidence from colombia. *Journal of International Economics*, 137:103598.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). Agricultural Productivity Differences across Countries. *American Economic Review*, 104(5):165–170.

Gutiérrez Sanín, F. (2019). Laberintos Institucionales: Una mirada crítica a los programas de formalización de la propiedad rural en Colombia. *Observatorio de Restitución y Regulación de Derechos de Propiedad Agraria*.

Ibánez, A., Munoz-Mora, J., and Gáfaro, M. (2012). Atlas de la distribution de la propiedad rural en colombia. *Universidad de los Andes*.

Ibáñez, A. M., Quigua, J., Romero, J., and Velásquez, A. (2022). Responses to temperature shocks: Labor markets and migration decisions in el salvador. *IDB Working Paper Series*.

Ibáñez, A. M. and Muñoz, J. C. (2010). The Persistence of Land Concentration in Colombia: What Happened Between 2000 and 2009? In Bergsmo, M., editor, *Distrubutive justice in transitions*, number 6 in FICHL publication series. Torkel Opsahl Acad. EPubl, Oslo.

IPCC (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. *Technical Summary*.

Jagnani, M., Barrett, C. B., Liu, Y., and You, L. (2021). Within-season producer response to warmer temperatures: Defensive investments by kenyan farmers. *The Economic Journal*, 131(633):392–419.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of political Economy*, 114(3):538–575.

Jessoe, K., Manning, D. T., and Taylor, J. E. (2018). Climate change and labour allocation in rural mexico: Evidence from annual fluctuations in weather. *The Economic Journal*, 128(608):230–261.

Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review*, 109(10):3585–3616.

Kazianga, H. and Udry, C. (2006). Consumption smoothing? livestock, insurance and drought in rural burkina faso. *Journal of Development economics*, 79(2):413–446.

Krusell, P. and Smith, Jr, A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy*, 106(5):867–896.

Madhok, R., Noack, F., Mobarak, A. M., and Deschenes, O. (2022). Internal migration and the re-organization of agriculture. *Working Paper*.

Mahul, O. and Stutley, C. J. (2008). Government support to agricultural insurance. *Challenges and options for developing countries. Washington, DC: The World Bank*.

Martinez, L. R. (2019). Sources of revenue and government performance: evidence from colombia. *Available at SSRN 3273001*.

Musyoka, P. K., Onjala, J., and Mureithi, L. P. (2021). Determinants of distress sales of farmland in rural kenya. *Development Studies Research*, 8(1):317–345.

Rao, M., Eberhard, J., and Bharadwaj, P. (2022). Towns and rural land inequality in india. *Working Paper*.

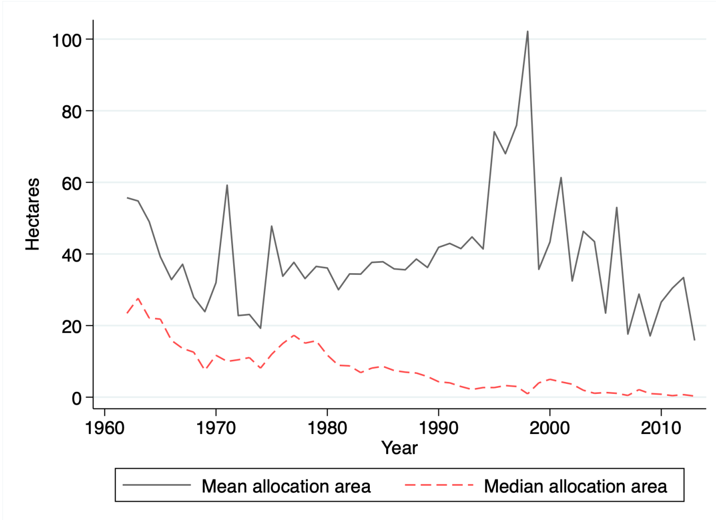
Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of monetary economics*, 55(2):234–250.

Rosenzweig, M. R. and Wolpin, K. I. (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in india. *Journal of political economy*, 101(2):223–244.

Sánchez, F. and Villaveces, M.-J. (2016). Tendencias y factores económicos y sociales asociados a la adjudicación de baldíos en Colombia, 1961-2010. In Cano, C. G., Iregui, A. M., Ramírez, M. T., and Tribín, A. M., editors, *El Desarrollo Equitativo, Competitivo y Sostenible del Sector Agropecuario en Colombia*. Banco de la República de Colombia.

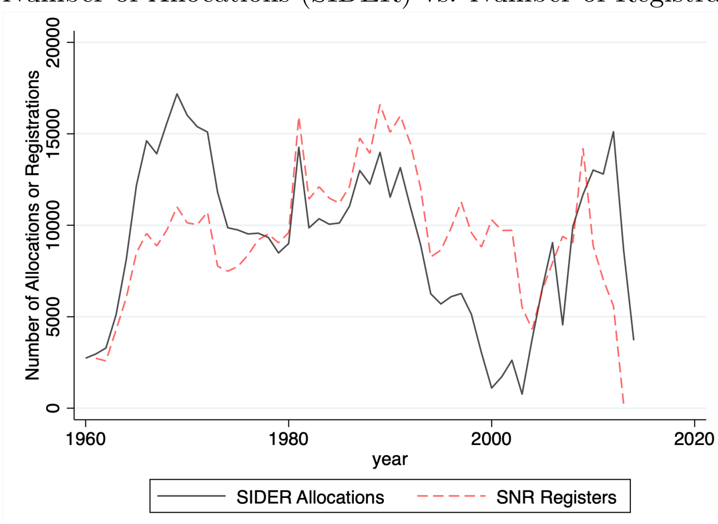
2.7 Appendix A: Additional Figures and Tables

Figure 2.6: Mean and Median Allocation Size - 1961–2012



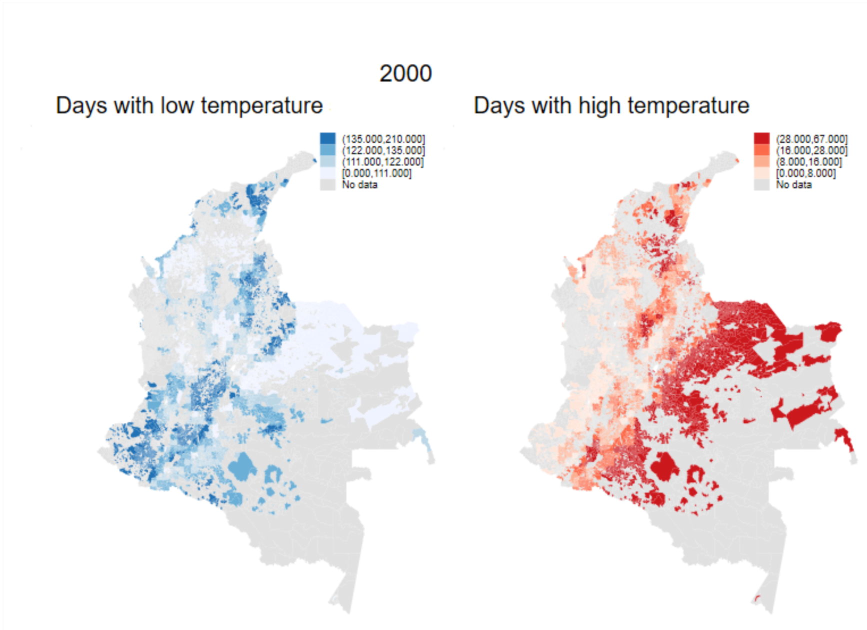
Notes: Data from the System of Information for Rural Development (SIDER). National-level yearly average area of land plots granted by the government as part of the public-land allocation program.

Figure 2.7: Number of Allocations (SIDER) vs. Number of Registrations (SNR)

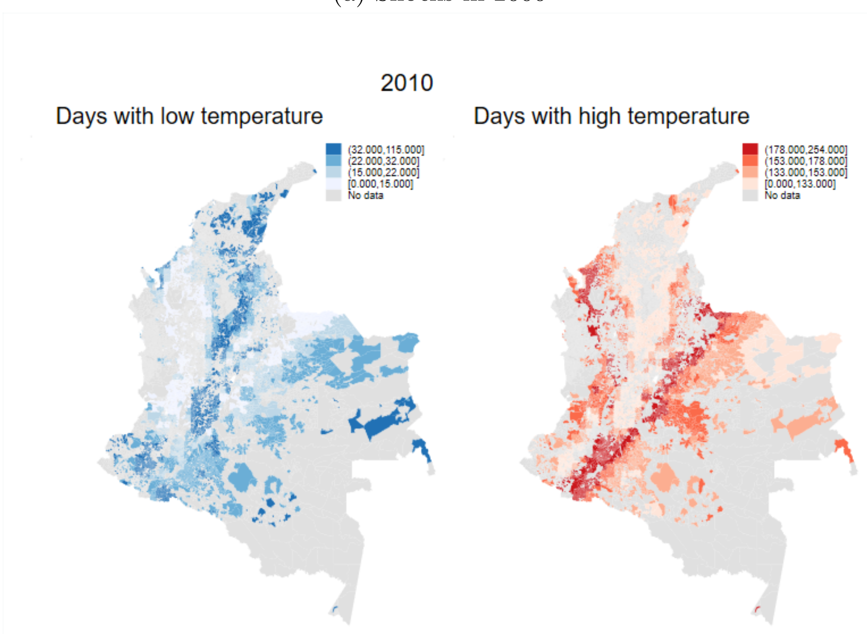


Notes: Data from the System of Information for Rural Development (SIDER) and from the National Superintendency of Notaries (SNR). The figure compares the number of land plots allocated by the government as part of the public-land allocation program with the number of properties registered at local public notary offices as received by the government. Property registration constitutes the final step to finalize the allocation process and ensures the formal property right of the beneficiary over the granted plot of land.

Figure 2.8: Temperature Shocks Across Space - 2000 and 2010



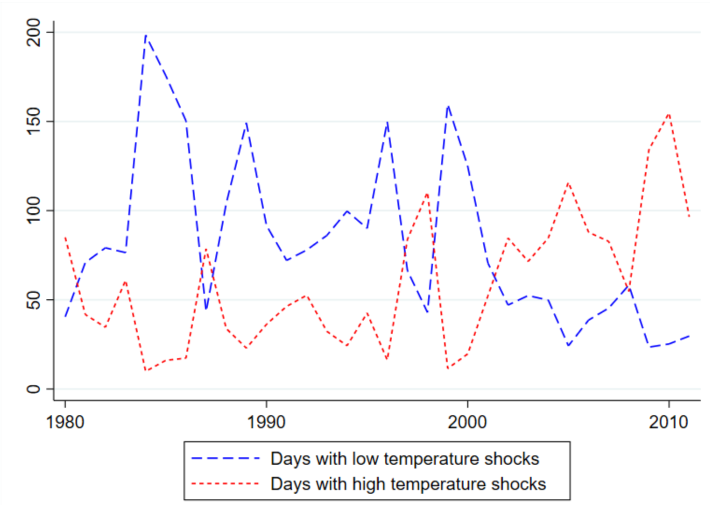
(a) Shocks in 2000



(b) Shocks in 2010

Notes: Data from the Copernicus Climate Change Service (C3S). The figure shows the average number of days with extreme heat (red) and cold (blue) across veredas in our sample in 2000 and 2010.

Figure 2.9: Temperature Shocks Across Time



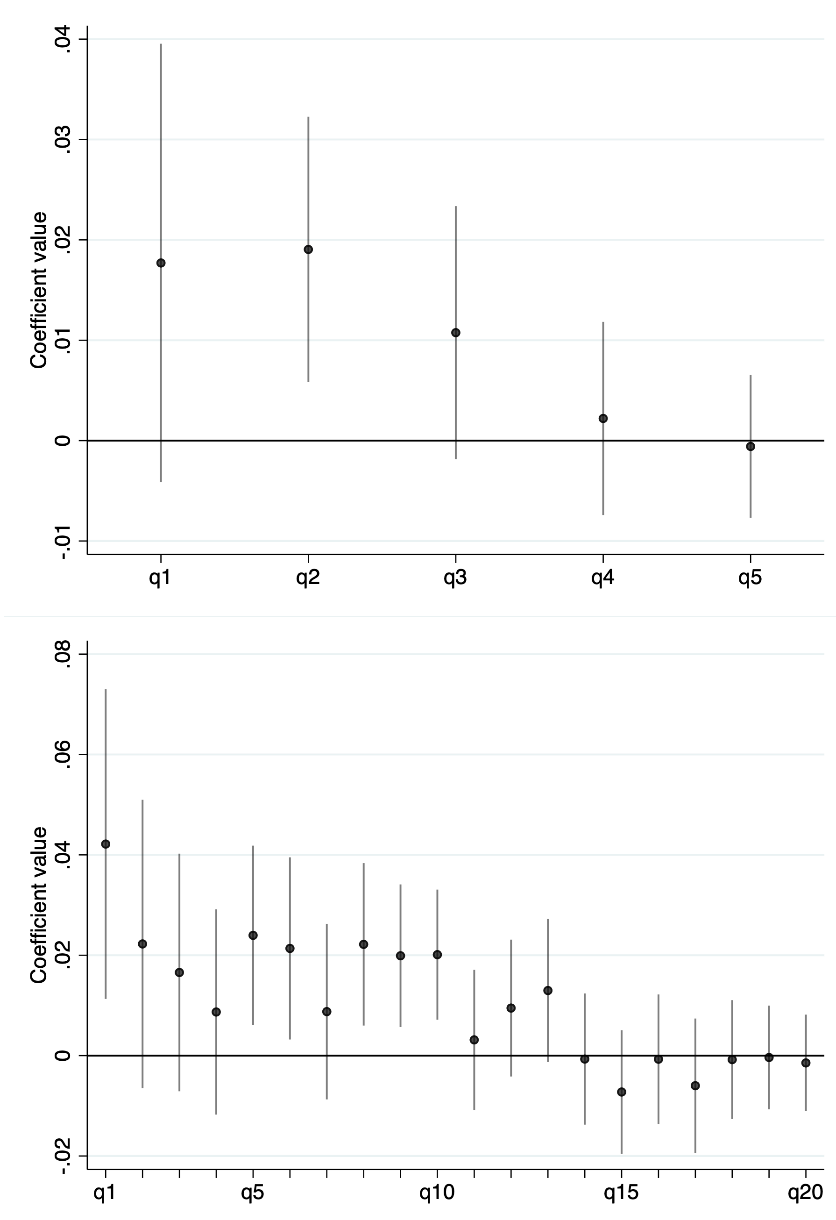
Notes: Data from the Copernicus Climate Change Service (C3S). The figure shows the average number of days with extreme heat (red) and cold (blue) across veredas in our sample for the 1979–2016 period.

Table 2.7: Temperature Shocks and Number of owners, by Initial Size Quantile

	Number of owners by initial distribution quantiles (q_m^j)									
	(1) q_m^1	(2) q_m^2	(3) q_m^3	(4) q_m^4	(5) q_m^5	(6) q_m^6	(7) q_m^7	(8) q_m^8	(9) q_m^9	(10) q_m^{10}
$TempShocks_{v,y}$	0.029*	0.016	0.024***	0.015**	0.020***	0.005	0.005	-0.004	-0.005	-0.000
	(0.015)	(0.010)	(0.008)	(0.008)	(0.006)	(0.007)	(0.006)	(0.005)	(0.006)	(0.004)
Observations	10915	10878	10853	10804	10907	10869	10928	10892	10907	10928
R^2	0.942	0.971	0.982	0.983	0.987	0.986	0.986	0.991	0.990	0.993

Notes: Data from the National Land Registry (*Catastro Nacional*), mantained by the National Geographical Institute (IGAC). Dependent variables are number of owners whose farm are is in the corresponding size range defined by the quantiles of the initial farm distribution. Dependent variables are in logarithms. The main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions also include year and geographic fixed effects (vereda or municipality). *Mean Dep. Var.* is the mean of the untransformed variable. Standard errors clustered at the municipality level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Figure 2.10: Temperature Shocks and Number of Owners by Initial Distribution Quantiles - Alternative Partitions



Notes: OLS estimates of the γ coefficients according to equation (2.5). Top panel: coefficient values for 5 quantiles of the initial municipality-level distribution of farm sizes. Bottom panel: coefficient values for 20 quantiles of the initial municipality-level distribution of farm sizes. Each point estimate corresponds to a separate regression where the main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions include year and municipality fixed effects. Error bars display 95% confidence intervals for standard errors clustered at the municipality level.

Table 2.8: Temperature Shocks and Land Sales - Alternative Specifications

	Municipality level panel				Vereda level panel			
	Total	Full	Partial	Mortg.	Total	Full	Partial	Mortg.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: State Specific Time Trends							
$TempShocks_{v,y}$	0.076***	0.087***	0.115***	0.103***	0.020***	0.022***	0.003	0.022***
	(0.021)	(0.023)	(0.028)	(0.020)	(0.006)	(0.006)	(0.005)	(0.006)
Observations	10,392	10,392	10,392	10,392	149,652	149,652	149,652	149,652
R-squared	0.912	0.903	0.710	0.794	0.574	0.562	0.360	0.393
Mean Dep. Var	1.54	1.44	0.44	0.64	0.24	0.22	0.04	0.06
	Panel B: Temperature Shocks with Absolute Thresholds							
$TempShocksAbs_{v,y}$	0.070**	0.108***	0.001	0.004	0.015	0.021***	-0.005	0.016*
	(0.028)	(0.029)	(0.029)	(0.031)	(0.009)	(0.008)	(0.007)	(0.009)
Observations	10,392	10,392	10,392	10,392	149,652	149,652	149,652	149,652
R-squared	0.912	0.903	0.709	0.793	0.574	0.562	0.360	0.392
Mean Dep. Var	1.54	1.44	0.44	0.64	0.24	0.22	0.04	0.06

Notes: Data from the National Superintendency of Notaries (SNR) records. Columns 1 and 5 show the effect on total (full + partial land sales) columns 2 and 6 show the effect on full sales (when the entire property is transferred to another owner), columns 3 and 7 show the effect on partial sales (when only a fraction of the plot is transferred), and columns 4 and 8 show the effect on mortgage originations. All dependent variables are in log(x+1) transformation. The main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions also include year and geographic fixed effects (vereda or municipality). Panel A adds state specific time trends. Panel B uses country level absolute thresholds to identify days of atypical temperature. *Mean Dep. Var.* is the mean of the untransformed variable. Standard errors clustered at the municipality level reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Table 2.9: Temperature Shocks and Average Farm Size - Alternative Specifications

	Number of Plots	Numbers of Owners	Mean Plot Size	Mean Area/Owner	Median Plot Size	Median Area/Owner
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: State Specific Time Trends					
$TempShocks_{v,y}$	0.012**	0.013***	-0.012**	-0.013***	-0.017	-0.013
	(0.005)	(0.005)	(0.005)	(0.005)	(0.012)	(0.009)
Observations	10,935	10,935	10,935	10,935	10,935	10,935
R-squared	0.990	0.992	0.994	0.995	0.976	0.988
Mean Dep. Var	2,518.49	2,516.07	30.49	29.36	15.22	12.88
	Panel B: Temperature Shocks with Absolute Thresholds					
$TempShocksAbs_{v,y}$	0.023**	0.016*	-0.023**	-0.020**	-0.032*	-0.012
	(0.009)	(0.009)	(0.009)	(0.009)	(0.016)	(0.012)
Observations	10,935	10,935	10,935	10,935	10,935	10,935
R-squared	0.990	0.992	0.994	0.995	0.976	0.988
Mean Dep. Var	2,518.49	2,516.07	30.49	29.36	15.22	12.88

Notes: Data from the National Land Registry (*Catastro Nacional*), mantained by the National Geographical Institute (IGAC). All dependent variables are in logarithms. The main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions also include year and geographic fixed effects (vereda or municipality). Panel A adds state specific time trends. Panel B uses country level absolute thresholds to identify days of atypical temperature. *Mean Dep. Var.* is the mean of the untransformed variable. Standard errors clustered at the municipality level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

2.8 Appendix B: The Land Ceiling Regulation

Finally, we investigate if our results on the absence of land consolidation on the part of large landholders in the aftermath of an adverse weather shock is due to institutional factors stemming from Colombia's land regulation policies. As discussed in section 3.2, Law 160 of 1994 imposed municipality-specific land ceilings that place a cap on the amount of land originally granted by the government that any private individual can accumulate. This restriction could be consistent explanation for the lack of land consolidation on the right part of the farm size distribution, since it restricts the capacity of large landholders to acquire any new land plots whose provenance was a government allocation.[15]

To test if these restrictions are in fact explaining our results, we re-estimate the model in (2.3) including an additional interaction term between the shock variable and a dummy indicating if the municipality is above the median in the share of the municipality's area that was at some point part of a government allocation. The idea behind this test lies in the fact that land ceilings only apply to allocated land, but not to other land plots. Hence, if restrictions are driving the land-fragmentation results shown in Table 2.3 we would expect the bulk of the result to be concentrated in municipalities with a high share of their agricultural land coming from government allocations.

As columns 5-8 in Table 2.10 show, we find no such heterogeneity. Moreover, as shown in columns 1-4, including the continuous value of the share of government-allocated land as a control has virtually no impact on the magnitude or precision of the original estimates. We take these results as evidence that the main findings of our paper are not driven by the specific institutional characteristics of land regulation in Colombia.

[15]The explicit purpose of the land ceilings, as stated in the text of the law, was precisely to prevent land concentration by large landholders.

Table 2.10: Temperature Shocks, Farm Size, and Share of Government-Allocated Area

	Control: Share Allocated				H_i: Share Allocated			
	Number of Farms	Number of Owners	Mean Farm Size	Mean Area/Owner	Number of Farms	Number of Owners	Mean Farm Size	Mean Area/Owner
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$TempShocks_{v,y}$	0.0113**	0.0112**	-0.0113**	-0.0115**	0.0134***	0.0116**	-0.0134***	-0.0119**
	(0.0049)	(0.0046)	(0.0049)	(0.0047)	(0.0048)	(0.0046)	(0.0048)	(0.0046)
$TempShocks_{v,y} \times H_i$					-0.0068	-0.0013	0.0068	0.0012
					(0.0092)	(0.0080)	(0.0092)	(0.0080)
Observations	10,934	10,934	10,934	10,934	10,935	10,935	10,935	10,935
R-squared	0.9905	0.9920	0.9935	0.9947	0.9905	0.9921	0.9935	0.9948
mean.dep.var	2519	2516	30.50	29.36	2518	2516	30.49	29.36
Share alloc.	Yes	Yes	Yes	Yes	No	No	No	No

Notes: Data from the National Land Registry (*Catastro Nacional*), maintained by the National Geographical Institute (IGAC). All dependent variables are in logarithms. The main independent variable is the total number of atypical temperature days in the past two years ($y-1$, $y-2$) divided by 100. Controls are accumulated allocations, accumulated precipitation during years y, $y-1$, and $y-2$. Regressions also include year and geographic fixed effects (vereda or municipality). *Mean Dep. Var.* is the mean of the untransformed variable. Standard errors clustered at the municipality level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Chapter 3

Mexican Migration Flows and Agricultural Labor Markets in the U.S.

With Ashish Shenoy

3.1 Introduction

Relative to population size, immigrant workers play an outsized role in several industries of the U.S. economy. In particular, Mexican-born migrants comprise almost 70% of hired agricultural workers in the country (Hernandez and Gabbard, 2019). However, migration from Mexico to the United States has steadily declined over the past twenty years, and the U.S. farm labor supply has been contracting for more than a decade. The prevalence of labor shortages in the agricultural industry is increasing (Zahniser et al., 2018), and the scarcity of available workers has intensified in the last few years (Peri and Zaiour, 2022). A 2019 survey found that 56% of farms in California reported being unable to meet their full labor demand over the previous five years. Common responses to the shortage included raising of wages, adopting labor-saving technologies, and reducing or delaying various cultivation practices.[1] As summarized by Martin (2020), farm employers must face lower migration flows through some combination of satisfying, stretching, substituting, or supplementing

[1] California Farm Bureau Federation. 2019. “Still Searching for Solutions: Adapting to Farm Worker Scarcity Survey”.

the available workforce. While conceptually any of these strategies seems a viable response to dwindling migration inflows, empirical evidence on the way in which agricultural producers have been adapting to these secular changes in labor availability remains scarce. Understanding the type of adaptation process that is taking place, quantifying the relative importance of potential adjustment margins, and—crucially—knowing how short-run responses differ from more structural adjustments in production practices, is necessary for the correct design of agricultural and migratory policy.

This paper estimates the causal effect of migration inflows on local labor market outcomes in the U.S. agricultural industry. Estimating how changes in worker arrivals affect local labor markets in agriculture has so far remained challenging given that a large fraction of the migrant agricultural workforce is undocumented. Reliable data on this type of migratory flows at a sufficiently disaggregated geographical level has usually been unavailable or lacking in representativity. We overcome this obstacle by using a novel, high quality administrative dataset based on consular identity cards issued by the Mexican government to (predominantly undocumented) workers, and in this way are able to measure the strength of migratory flows between every Mexican municipality and U.S. county pair throughout the 2008–2019 period.

We instrument for migration flows of Mexican workers into U.S. counties by using a shift-share design that combines the sudden and spatially heterogeneous increase in violence across Mexican municipalities beginning in 2008 with preexisting migration networks. Violence-driven migration flows in origin municipalities are plausibly orthogonal to labor demand conditions in destination counties, while location choices are well predicted by the share of migrants from the same municipality already established in each potential destination before violence levels rose. Based on this identification strategy, and combining migration flows data with county-level information on agricultural wages, employment, and cultivation practices, we are able to quantify how variations in the supply of migrant workers shape employers' hiring and production decisions across different agricultural sub-industries.

Our estimates reveal a stark contrast between the short-run and the long-run effects that migratory flows have on agricultural labor markets. In the short run, we find that yearly reductions in migration rates from Mexico put upward pressure on wages across all types of agricultural workers, with a one percentage point decrease in migration rates leading to wage increases ranging between 0.22% and 0.42% depending on the specific sub-industry analyzed. These magnitudes are in line with other estimates of the response of wages to migration in contexts where incumbent workers and new arrivals appear to be relatively close substitutes (Kleemans and Magruder, 2018; Imbert et al., 2022).

The rise in wages produced by lower migration is accompanied by a reduction in the number of workers directly hired by producers. This reduction, however, is more than offset by a large increase in the number of H-2A seasonal worker visa requests made by producers, with a one percentage point reduction in migration rates simultaneously causing a reduction of 0.6% in directly-hired employment and an *increase* in H-2A visa requests of 3.4%. To the best of our knowledge, our findings provide the first causal estimate of the elasticity of substitution between permanent immigrant worker arrivals and H-2A temporary visa requests, and confirm the hypothesis that increasing the demand for guest workers is the main margin of adjustment through which U.S. producers adapt, in the short run, to reductions in the labor supply.

By contrast, our estimates for the long-run response to migration flows show that U.S. counties experiencing more severe reductions in migrant-labor supply during the 2008–2019 period had *lower* average wage growth by the end of the period, accompanied by slightly lower levels of agricultural employment. Our estimates indicate that a one percentage point reduction in the cumulative–annualized—migration rate between 2008 and 2019 caused a decrease in average weekly wages ranging between 0.15% to 0.27% across different worker types.

Why do higher immigration rates lead to higher wages? The differences between short and long run responses to migration suggest that both labor markets and the broad agricultural production process is adjusting other less flexible factors of production like land and capital as time goes by. Lower expected labor availability might trigger mechanization processes that substitute production from labor to machinery as documented by (Clemens et al., 2018). Another explanation could rely on the existence of complementarities between domestic and foreign workers (Ottaviano and Peri, 2012), where increases in the supply of foreign workers spurs productivity increases for incumbent workers leading to higher wages. Higher migration rates might lead to higher long-run wages in a sector if, for example, a larger pool of potential employees allows producers to select more skilled workers that lead to productivity increases. It is also possible that increases in the supply of available labor might spur economic activity in other industries of the economy where a large share of workers are also foreign immigrants (Charlton and Castillo, 2022), with this increased competition for migrant labor across sectors leading to faster wage growth. Finally, the tradable nature of agricultural goods might also imply that a higher supply of workers is eventually adjusted for mostly through an expansion in output and revenues rather than through lower wages or higher unemployment (Burstein et al., 2020).

We explore these potential mechanisms and investigate how long-run variations in migration rates change production decisions by comparing county-level changes in agricultural practices as measured in the Census of Agriculture. This analysis reveals that counties with exogenously fewer migrant arrivals between 2007 and 2017 shifted their crop composition away from labor-intensive crops like vegetables, fruits and tree nuts, and increase instead the area devoted to field crops. We further find that lower migration rates led to relative increases in mechanization, measured both as the change in the absolute value of machinery employed and also in proportion to land area cultivated and number of workers employed. Consistent with migration-driven increases in productivity, our estimates show that higher migration rates cause average farmland values in a county to increase, with a one percentage point increase in annual migration causing a 3.4% increase in farmland values measured in dollars per acre. Finally, we find that higher migration rates do not lead to increases in total output or sales, but rather seem to lead to lower average incomes reported by producers.

This paper contributes to the broad literature on the labor market impacts of immigration (Card, 2001; Borjas, 2003), highlighting how the short run responses to migration flows hinge on the degree of substitutability between incumbent and arriving workers, and by showing how in the long run these effects can be attenuated, or even reversed, through adjustments in other margins. The paper also contributes to the literature focused on studying the causes and consequences of migration from Mexico to the U.S. (Hanson and McIntosh, 2010), and in particular to the effects that the sustained slowdown in migration rates can have in the productive potential of the agricultural sector (Charlton and Taylor, 2016; Rutledge and Mérel, 2023).

The rest of the paper is organized as follows. In Section 2 we describe the general trends on migratory flows from Mexico and agricultural labor, present the data and describe the samples and variables used for analysis. In Section 3 we motivate the use of our shift-share instrument, and discuss our empirical methodology. Section 4 reports our empirical results, and Section 5 concludes.

3.2 Data and Background

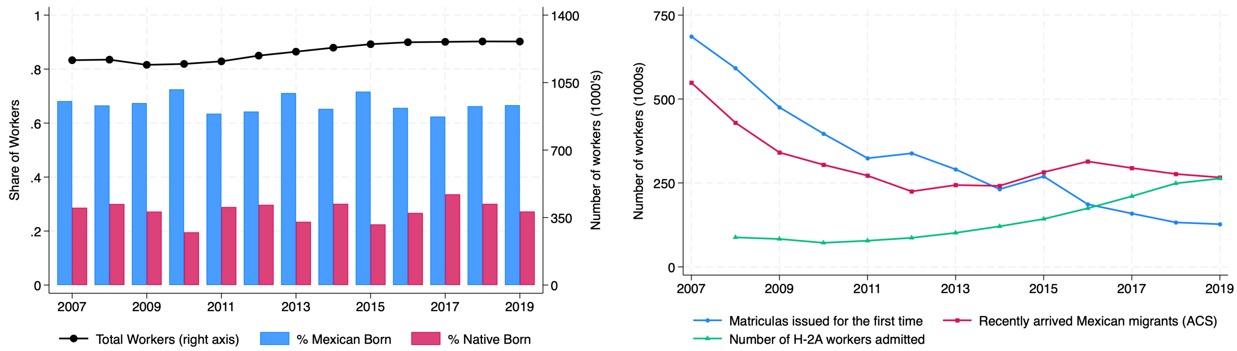
3.2.1 Mexican migration and agricultural labor: National-level trends

Our findings coincide with national-level trends showing that, while both the total number of workers and the share of Mexican-born workers hired in agriculture have remained stable during the 2000–2020 period, there has been a substantial drop in immigration flows arriving to the U.S. from Mexico. The sustained decline in worker arrivals from Mexico has been large enough that, at least since 2014, the net flow of migrants from Mexico to the U.S. is roughly zero and might even be slightly negative (Gonzalez-Barrera, 2015). By contrast, as Figure 3.1 illustrates, the number of temporary H-2A workers being requested by U.S. farmers in order to fill vacant positions increased from under 87,000 yearly workers in 2008, to more than 250,000 in 2019. Taken together, these trends suggest that—against the backdrop of falling permanent immigrant arrivals–U.S. agricultural producers are increasingly relying on H-2A guest workers as a source of labor.

3.2.2 Migratory flows: the *Matriculas Consulares de Alta Seguridad* data

We measure the yearly inflow of Mexican workers arriving to each U.S. county using information from the *Matrículas Consulares de Alta Seguridad* (MCAS) program maintained by the Mexican government. The MCAS dataset records all *Matrículas Consulares* identification cards issued by Mexican consulates to Mexican-born individuals living in the U.S., and registers both the municipality of origin and the current U.S. county of residence of each cardholder. MCAS are issued to all qualifying Mexican citizens regardless of immigration status or age. While holding a

Figure 3.1: U.S. Agricultural Workers and Mexican Migration Inflows - 2007–2019

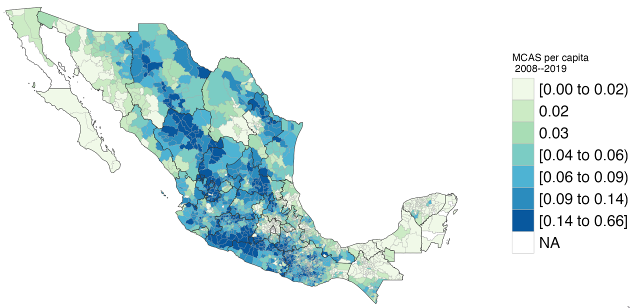


Notes: National-level trends on agricultural sector workers and immigration flows from Mexico to the U.S. Left panel: Total number of hired workers in the agricultural sector according to the Quarterly Census of Employment and Wages (*QCEW*), and share of hired workers who are either Mexican-born or native born according to the National Agricultural Workers' Survey (*NAWS*). Right panel: Number of Matriculas Consulares de Alta Seguridad (*MCAS*) issued for the first time by all Mexican consulates in the U.S. (blue line); number of workers born in Mexico arrived to the U.S. within the previous year according to the American Community Survey (*ACS*) (red line); number of H-2A guest workers admitted by the Department of Labor (green line).

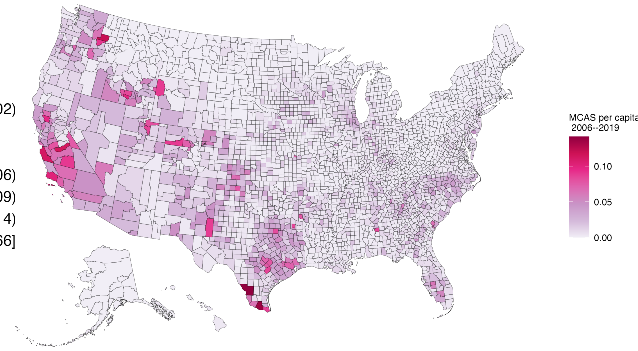
MCAS card does not confer any U.S. immigration status to the person to whom it is issued, many states and local governments allow the document to be used as proof of identity, and so it permits the cardholder to access a services that include opening a bank account, being assigned an Individual Tax Identfication Number (ITIN), or obtaining a driver's license. Since obtaining the MCAS does not entail any additional benefit to authorized migrants, it is generally assumed that MCAS are a measure of unauthorized migration inflows to the U.S. (Massey et al., 2010). Comparing the joint distribution of these inflows both at origin and destination with alternative surveys, Caballero et al. (2018) confirm that MCAS records are in fact a representative and high-quality information source on Mexican migratory flows.

We take yearly MCAS issued for the first time as our main measure of Mexican migration flows to the U.S. from Mexico. Given that MCAS must be renewed every five years it is important, when trying to measure yearly migrant inflows, to separate renewals from first-issuances. The right panel of Figure 3.1 shows the close correspondence between the observed number of MCAS issued for the first time and the number of newly-arrived Mexican migrants to the U.S. as recorded in the *American Community Survey* (ACS), and is consistent with the documented decline in migration inflows throughout this period (Passel and Cohn, 2018). Figure 3.4 in Appendix 3.6 further shows the correspondence between ACS and MCAS data for the four states in the country with highest

Figure 3.2: Migration rates in MCAS data – Municipality-level outflows and County-level inflows



(a) Municipality outflows per capita



(b) County inflows per capita

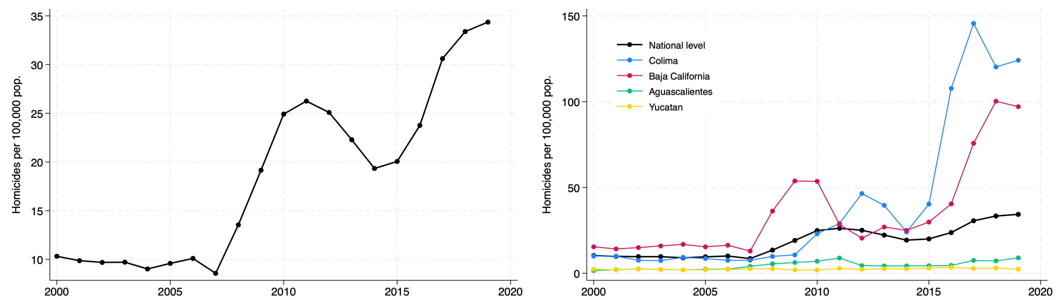
immigration rates. A detailed analysis of the validity of MCAS data as a measure of migration flows across time can be found in Tiburcio and Camarena (2023). The spatial distribution of municipality-level outflows, and county-level inflows of migrants is shown in Figure 3.2.

3.2.3 Mexican migration networks

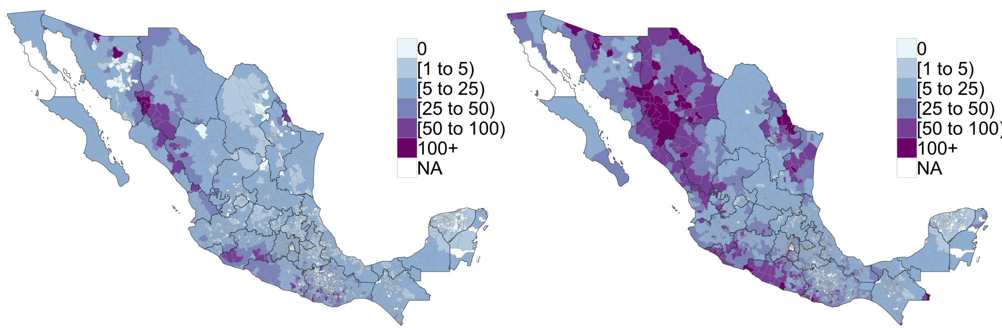
The MCAS records allow us to build a measure of the strength of migratory networks between each Mexican municipality of origin and destination U.S. county pair. A large literature (see for example Munshi (2003, 2014)) shows that preexisting migrant community networks at destination are a fundamental determinant of the location choice of future migrants. To compute this network measure we use the first two years in our data (i.e. 2006 and 2007) and calculate the share of migrants going to each U.S. county out of the total number of migrants leaving each Mexican municipality during this two-year period. As documented by Tian et al. (2022), we find large differences in the historical destination patterns across different Mexican municipalities, even within the same state. We leverage this spatial variation in the settlement patterns of different origin communities to build the shares of our shift-share instrument for migration inflows. Figure 3.5 in Appendix 3.6 illustrates the differences in migration networks for a pair of nearby municipalities.

Figure 3.3: Homicide Rates in Mexico – 2000–2019

(a) National homicide rate 2000-2019 (b) Homicide rates - selected States



(c) Municipal homicide rates 2000–2007 (d) Municipal homicide rates 2008–2019



Notes: Yearly homicide rates per 100,000 population. Source: Mexican national statistics institute (INEGI).

3.2.4 Violence in Mexico

Starting in 2008, Mexico has suffered an unprecedented and drastic increase in violence levels across large parts of its territory. As shown in Figure 3.3, national homicide rates increased by 300% in 12 years, from 8.6 homicides per 100,000 population in 2007 to 34.4 homicides per 100,000 population in 2019 . While most violence in Mexico is closely related to the illicit drug market supply to the U.S. and the governmental stance on the war on drugs, the specific causes of the sudden surge in homicide rates remain a source of debate (Castillo et al., 2020; Guerrero-Gutiérrez, 2011; Dell, 2015; Williams, 2012). Regardless of their cause, these rapid increases in violence levels across the country have been shown to have negative impacts on labor-market outcomes (Velásquez, 2020), capital accumulation (Brown and Velásquez, 2017), and migration decisions (Orozco-Aleman and Gonzalez-Lozano, 2018).

We use yearly municipal-level data on violent homicide rates from INEGI, Mexico's national statistical agency. As Figure 3.3 shows, the growth in violence levels has been heterogeneous across space, with some states suffering a more-than-tenfold increase in murder rates while others showing no increase at all throughout the period. Combining this data with the MCAS information described above, in Section 3.3 we document the existence of a strong positive correlation between yearly municipal violence levels and migrant outflows to the U.S.

The spatial variation in the increase in violence across the 2008–2019 period acts as the 'shift' component of our shift-share instrument. By combining this component with municipal-county network shares we are able to aggregate destination origin shocks into a yearly destination-level measure of migrant-supply shocks.

3.2.5 Agricultural wages, employment, and practices

To measure labor market outcomes we use county-industry-year level data on wages and employment from the *Quarterly Census of Employment and Wages* (QCEW) program mantained by the U.S. bureau of labor statistics. The QCEW publishes a quarterly count of establishments, employment level, and total wage bill at the 6-digit NAICS industry level for each county in the United States. It is based on the aggregation of all Quarterly Contribution Reports (QCRs) submitted by employers subject to state and federal unemployment insurance laws to each State's accounting system. The QCEW covers more than 95% OF U.S. jobs. We focus on employment levels and wages recorded in the QCEW under NAICS codes 11 (agriculture, forestry, fishing and hunting), 111 (crop production), and 115 (Agriculture and forestry support activities). Figure 3.6 in Appendix 3.6 show national-level trends of employment and wages for these sub-industries. Agricultural labor tends to be highly seasonal, and about half of all hired workers tend to be hired directly and employed on crop production. Indirectly-hired workers in agriculture support activities make up roughly 20% of the total workforce and, until recently, tended to have lower earnings than their directly-hired counterparts.

We also estimate the impact of migration on the yearly rates of H-2A workers requested for authorization by agricultural employers in each U.S. county. Data on these requests are publicly available at the individual request level from the Department of Labor. [2] To get yearly county-

[2] https://www.dol.gov/agencies/eta/foreign-labor/performance

level requests we aggregate the total unique number of workers certified on each intended worksite according to the stated job start date on each application.[3] Figure 3.1 shows yearly trends in national H-2A request levels, and Figure 3.9 in Appendix 3.6 shows the change in the distribution of H-2A requests across space. H-2A requests are concentrated in the southern states of the U.S.—Louisiana, Florida, North Carolina—, as well as in California, Colorado and the state of Washington. Since 2008, the number of certified H-2A seasonal workers has more than doubled from under 100,000 workers to nearly 250,000 workers in 2018.

We use data from the 2007 and 2017 versions of the Census of Agriculture (COA) to measure the effect of migration on agricultural practices. The Census of Agriculture, carried out every 5 years, is a complete count of all farms and ranches in the country and the people who operate them. It collects information on land use and ownership, operator characteristics, production practices, income and expenditures.

3.3 Methodology

3.3.1 Instrumental variable motivation

Assessing whether changes in migration rates have an effect on wages or employment rates is challenging due to the fact that observed wages and employment are equilibrium outcomes that are endogenous to labor supply, of which migration is only one component, and to labor demand, which is affected by the local economic cycle. To overcome this challenge we use the interaction of changes in violence levels across Mexican municipalities with observed preexisting migration networks as an instrument for the number of migrants arriving to each destination county every year. This instrument is based on the observation that variations in the intensity of violence in origin locations are drivers of the decision to migrate, and that the destination choice is further driven by the strength of social networks created by previous migration waves.

To fix ideas, let $M_{c,t}$ be the number of migrants arriving to U.S. county c on year t. (i.e. the number of first-time issued MCAS assigned to county c). Our goal is to find an instrument for the

[3] In some years intended worksites are specified either as cities or zip codes. We harmonize all locations at the U.S. county level using the crosswalks provided by the Missouri Census Data Center https://mcdc.missouri.edu/applications/geocorr2018.html.

yearly Mexican immigration rate to each U.S. county c:

$$m_{c,t} = \frac{M_{c,t}}{P_{c,t^0}}$$

while the emigration rate leaving to the U.S. from municipality o on year t is

$$n_{o,t} = \frac{M_{o,t}}{P_{o,t^0}}$$

where $P_{o,t}$ and $P_{c,t}$ are, respectively, municipality and county populations at year t and t^0 indicates a year prior to t.[4]

To test the premise that violence levels at origin municipalities are in fact correlated with migratory outflows we regress yearly municipality-level emigration rates $n_{o,t}$ on yearly homicide rates $V_{o,t} = \frac{\text{Homicides}_{o,t}}{P_{o,t^0}}$

$$n_{o,t} = \alpha + \beta V_{o,t} + \delta_t + \gamma_o + \varepsilon_{o,t} \tag{3.1}$$

where δ_t and γ_o are respectively year and municipality fixed effects.

Results for regression 3.1 are shown in Table 3.1. After accounting for both year and municipality fixed effects, our point estimate indicates that, on average, a 10 percentage point increment in the homicide rate is associated with a 2.8 percentage point increase in municipal emigration rates to the U.S. The magnitude of this correlation is similar to other estimates of the violence-at-origin effect on U.S. migration rates coming from other Central American countries (Clemens, 2021).

While we cannot tell if the observed association of violence and migrant outflows is causal, the existence of this strong correlation is enough motivation to use violence rates as the 'shift' component of our shift-share design. The intuition behind the second part of our instrument is that origin-destination migratory flows can be accurately predicted from aggregate municipality-level outflows multiplied by a measure of the strength of the historical settlement network between each municipality and each U.S. county. More precisely,

[4]In practice, we normalize all of our per capita variables according to the 2005 county and municipality population estimates calculated respectively by the U.S. Census and INEGI.

Table 3.1: Homicide rates and yearly emigration rates – Mexican municipality level

	Yearly emigration rate ($n_{o,t}$)			
	(1)	(2)	(3)	(4)
Homicides per capita	0.956***	1.293***	-0.213	0.283**
	(0.176)	(0.188)	(0.178)	(0.138)
Observations	29232	29232	29232	29232
Municipalities	2436	2436	2436	2436
Year FE	No	Yes	No	Yes
Municipality FE	No	No	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the municipality level in parentheses.

$$m_{c,t} = \frac{M_{c,t}}{P_{c,t^0}} = \frac{1}{P_{c,t^0}} \sum_o M_{o,c,t} \approx \frac{1}{P_{c,t^0}} \sum_o \left[M_{o,t} \times \phi_{o,c}^{t^0} \right] = \frac{1}{P_{c,t^0}} \sum_o \left[(n_{o,t} \times P_{mot^0}) \times \phi_{o,c} \right] \quad (3.2)$$

where $M_{o,c,t}$ is the migration flow from m to c in t, and the share of total migrants from municipality m that arrived to county c during the 2006–2007 two-year period:

$$\phi_{o,c} \equiv \frac{M_{o,c,t^0}}{\sum_c M_{o,c,t^0}}$$

is our measure of migrant-network strength. Whether this measure is indeed a good predictor of subsequent migrant location decisions can be evaluated in the data: Figure 3.10 shows that this is indeed the case when using the MCAS data, and that network-predicted migration flows (i.e. $[M_{o,t} \times \phi_{o,c}]$ in equation 3.2 are accurate predictors of observed county-municipality migration flows $M_{o,c,t}$.

Leveraging the fact that changes in municipal homicide rates influence emigration intensity, and that historical migration patterns are good predictors of destination choice, we construct the following shift-share instrumental variable:

$$Z_{c,t} = \frac{1}{P_{c,t^0}} \sum_o \left[\text{Homicides}_{o,t} \times \phi_{o,c}^{t^0} \right] \quad (3.3)$$

Table 3.2: County level immigration rates and origin violence shocks – First-stage estimates

	Yearly immigration rate ($m_{c,t}$)			
	(1)	(2)	(3)	(4)
$Z_{c,t}$: Violence shift-share IV	7.617***	7.862***	-5.318***	-4.272***
	(0.463)	(0.486)	(0.869)	(0.774)
Observations	37680	37680	37680	37680
Counties	3140	3140	3140	3140
Year FE	No	Yes	No	Yes
County FE	No	No	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the county level in parentheses.

3.3.2 First-stage results

To evaluate if the instrument is a strong predictor of county-level migrant inflows we run the following regression:

$$m_{c,t} = \alpha + \delta Z_{c,t} + \delta_t + \gamma_c + \varepsilon_{c,t} \tag{3.4}$$

where δ_t and γ_c are respectively year and county fixed effects. Results for regression equation 3.4 are shown in Table 3.2. While the simple pooled cross-section comparison of county-year observations shows a strong positive relationship between the instrument and migration rates, once county fixed effects are included and unobserved time-invariant county characteristics are accounted for, this relationship reverses and becomes strongly negative. This is a surprising result and could be subject to a number of different explanations. Additional estimations in appendix 3.7 show that this negative relationship is not due to *i)* The networks component ($\phi_{o,c}$) of the shift-share variable, *ii)* Noisiness of the yearly data, nor *iii)* The aggregation of violence measures across various municipalities.

We interpret the observed results as suggesting that, when comparing across counties, higher average violence levels in the group of municipalities associated to each county are strongly correlated to higher immigration rates, but that this migration tends to happen in relatively less-violent years. That is, while the yearly comparison across counties shows there is a clear positive relationship between violence and migration rates—i.e. counties with stronger connections to more violent municipalities have higher immigration rates—a within-county comparison yields that yearly deviations from county trend in aggregate municipal violence levels is negatively associated to migratory flows when compared to less violent years.

This interpretation is consistent with a model where violence-induced outmigration is determined by a long-run violence component that follows the migrant network decision rule, and by a short-run violence component that follows some independent short-run decision rule. Appendix 3.8 describes in more detail such a model, and shows that estimations based on the simplest version of this model—carried out on simulated data—are capable of replicating the change in sign of the first-stage regression coefficients observed after the inclusion of county fixed effects.

3.3.3 Empirical Strategy

To estimate the short-run effect of changes in migration to county c in year t we estimate a regression of the form:

$$y_{c,t} = \alpha + \beta^S m_{c,t} + X'_{c,t}\gamma + \tau_c + \delta_t + e_{c,t}, \tag{3.5}$$

where the dependent variable yc, t denotes some county-level outcome, $m_{c,t}$ is the immigration rate in county c at year t, and τ_c and δ_t are county and year fixed effects. The vector $X'_{c,t}$ includes the state-level minimum wage at year t and a Bartik-style shock that controls for time-varying changes to local labor demand.[5] Equation (3.5) is estimated through two-stage least squares using the instrumental variable defined in equation (3.3).

To estimate the long-run effect of migration inflows on labor markets we compute the county-level change in outcomes between 2008 and 2019 and regress it on the annualized sum of yearly migration flows relative to county baseline population for the same period:

$$\Delta y_c = \alpha + \beta^L \tilde{m}_c + X'_c\rho + v_c, \tag{3.6}$$

[5] The inclusion of these controls is meant to control for time-varying unobservable characteristics that might affect county-level migration inflows. However, excluding them from the regressions does not have any impact in the results for any outcome. The Bartik shock is computed from the Census *County Business Patterns* (CBP) data. It combines the 2-digit NAICS code industry composition of each county in 2007 with yearly national-level industry growth rates.

where Δy_c denotes the long difference in some county level outcome between 2008 and 2019, and $\tilde{m}_c \equiv \frac{1}{12}\sum_{t=2008}^{2019} m_{c,t}$, is the cumulative annualized migration rate. The vector X_c' is composed of controls for minimum-wage growth and long-differences in Bartik-style labor demand shocks. Migration rates are instrumented with the sum of the instrument across all periods $Z_c = \sum_{t=2008}^{2019} Z_{c,t}$. All standard errors in both short-run and long-run specifications are clustered at the county level.

Sample choice and regression weights In order to protect respondent's confidentiality, the published QCEW data suppresses information for industry-quarter-county combinations where the number of establishments is deemed small enough as to make individual information identifiable. This implies that the QCEW is, in practice, an unbalanced panel where industry-specific information for a given county in a given year might be undisclosed due to the small number of respondents. For each of the agricultural sub-industries mentioned above, our baseline estimation sample uses all counties that have complete information for all years in the analysis period for the specific sub-industry. Each regression in our baseline results is therefore carried out on a balanced panel of counties, but the set of counties can vary across sub-industry. To test for the sensitivity of our results, we estimate all main results using two alternative estimating samples that respectively consist of *i)* all county-year observations available, and *ii)* the 'fully restricted' sample of counties that have complete information for all industries in all years.

Similarly, our results could be dependent on the choice of weights specified in each estimation. While in our baseline results all counties in the estimating sample are given equal weights, we also test for the sensitivity of results relative to this assumption by estimating regressions weighted by *i)* baseline county population, and *ii)* baseline farm employment levels.[6] Results for these alternative specifications are shown in Figures 3.7 and 3.8 in Appendix 3.6. In general terms we find our results are robust to the choice of estimation sample and weighting scheme.[7]

[6]Baseline defined as population and employment in 2005.

[7]Results only differ qualitatively for a single outcome, long run H-2A employment, for the case of 'fully restricted' sample and for the case of baseline farm employment weights. Both of these specification give particular importance to counties with relatively large agricultural sectors, suggesting that, for counties highly specialized in agriculture, the long-run impact of higher migration rates on the prevalence of the H-2A worker program might be actually positive instead of null.

3.4 Results

3.4.1 Short-run Results

We estimate the impact of yearly migration flows on wages and employment for four different groups of workers: *i)* all workers hired in the agricultural sector (NAICS 11), *ii)* workers directly hired by employers for crop production (NAICS 111), *iii)* indirectly hired workers for crop support activities (NAICS 115), and *iv)* H-2A seasonal guest workers. Table 3.3 reports the OLS and IV estimates of equation (3.5) on the (log) average weekly wage for each of these worker types, while estimates for the effect on (log) employment levels are displayed on Table 3.4.

Table 3.3: Migration and Agricultural Wages – Short run effects

	(1) Total Agricultural	(2) Directly Hired	(3) Contract Labor	(4) H-2A Workers
Panel A: OLS Estimates				
Migration Rate ($m_{c,t}$)	-0.010	-0.064***	-0.028	-0.030
	(0.021)	(0.023)	(0.050)	(0.022)
N	11508	11808	4572	7848
Num Counties	959	984	381	654
Panel B: IV Estimates				
Migration Rate ($m_{c,t}$)	-0.221**	-0.325***	-0.323***	-0.416**
	(0.107)	(0.069)	(0.118)	(0.170)
N	11508	11808	4572	7848
Num Counties	959	984	381	654
Kleinberg-Paap F	10.49	17.4	7.28	9.955
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the county level in parentheses.

Panel A of Table 3.3 shows that variations in the number of Mexican migrants arriving to a county have in general a weakly negatively correlation with average agricultural wages. However, given that migration is endogenous to labor demand, and migrants are likely attracted to labor markets where wages are higher, the OLS estimates are likely underestimating the causal effect of migration on wages. Indeed, the IV estimates in Panel B of the same table show that increases in yearly migration rates have a much stronger negative effect on wages than the one suggested by the

OLS regressions. A one percentage point increase in migration rates causes a reduction of 0.22% in average weekly wages for all workers in the agricultural sector, and this effect intensifies when focusing both on the directly hired or contract labor workers. The largest reduction in average wages due to migration (0.41%) is on the wages paid to H-2A guest workers.

Table 3.4: Migration and Employment in Agriculture – Short run effects

	(1) Total Agricultural	(2) Directly Hired	(3) Contract Labor	(4) H-2A Workers
Panel A: OLS Estimates				
Migration Rate ($m_{c,t}$)	0.060	0.233***	-0.115	-0.515**
	(0.068)	(0.074)	(0.103)	(0.221)
N	11508	11808	4572	7908
Num Counties	959	984	381	659
Panel B: IV Estimates				
Migration Rate ($m_{c,t}$)	0.246	0.630***	-0.237	-3.418***
	(0.285)	(0.242)	(0.238)	(1.148)
N	11508	11808	4572	7908
Num Counties	959	984	381	659
Kleinberg-Paap F	10.49	17.4	7.28	10
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the county level in parentheses.

Similarly, the OLS estimates for the correlation between migration and employment shown in panel A of Table 3.4 underestimate the true effect. The IV estimates show that while a one percentage point decrease in yearly migration rates changes the number of workers hired directly by farmers by -0.63%, this reduction is more than compensated by the increase in the number of H-2A visa requests, which, at 3.4% is about five times larger. For its part, increases in migration also appear to have a negative effect with the number of indirectly-hired contract workers, but this relationship is not statistically distinguishable from zero.

Both sets of results suggest that, in the short run, agricultural labor demand in the U.S. is relatively inelastic, and that fluctuations in labor supply lead to wage responses across all type of farm workers as well as to large increases in the demand for foreign, seasonal guest workers. This is consistent with the seasonal characteristics of agricultural production, where production decisions

Table 3.5: Migration and Agricultural Wages – Long run effects

	(1) Total Agricultural	(2) Directly Hired	(3) Contract Labor	(4) H-2A Workers
Panel A: OLS Estimates				
Migration Rate (m_c)	0.130**	0.268***	0.232**	0.070
	(0.051)	(0.047)	(0.115)	(0.045)
N	959	984	381	654
Num Counties	959	984	381	654
Panel B: IV Estimates				
Migration Rate (m_c)	0.170***	0.235***	0.265**	0.157*
	(0.059)	(0.057)	(0.129)	(0.083)
N	959	984	381	654
Num Counties	959	984	381	654
Kleinberg-Paap F	144.98	130.02	48.21	131.96
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the county level in parentheses.

like crop choice and technology decisions have to be taken well in advance of labor recruitment. However, while yearly labor demand might be relatively fixed, long run adjustments on other margins of production might lead to changes in the elasticity of labor demand as shown below.

3.4.2 Long-run Results

Tables 3.5 and 3.6 present the results from estimating equation (3.6) on the log-difference of employment and wages between 2008 and 2019 for the same four categories of agricultural workers.

In contrast to short-run responses, these long-difference estimates reveal that, in the long run, increased migration rates induce *higher* average wage growth for all types of agricultural workers. In absolute value, the magnitude of the long-run wage responses to migration is smaller than the short-run response, with a once percentage point increase in annualized migration rates leading to average wage increases of between 0.16% and 0.26%. The weaker magnitude in wage response is consistent with the idea that adjustments in other factors of production across time allow labor demand to become less inelastic.

Table 3.6: Migration and Employment in Agriculture – Long run effects

	(1) Total Agricultural	(2) Directly Hired	(3) Contract Labor	(4) H-2A Workers
Panel A: OLS Estimates				
Migration Rate (m_c)	-0.154	-0.571***	0.257	1.105*
	(0.161)	(0.168)	(0.234)	(0.585)
N	959	984	381	659
Num Counties	959	984	381	659
Panel B: IV Estimates				
Migration Rate (m_c)	0.072	-0.231	0.420*	0.088
	(0.249)	(0.317)	(0.251)	(0.778)
N	959	984	381	659
Num Counties	959	984	381	659
Kleinberg-Paap F	144.98	130.02	48.21	133.843
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the county level in parentheses.

Regarding employment, the long difference estimates show no effect of higher migration rates on overall employment rates in agriculture or in the number of seasonal agricultural workers requested. We find increased migration leads to a relatively small increase in the number of indirectly-hired contract workers. The IV estimates stand in contrast with the OLS coefficients, which show a positive correlation between migration rates the log difference in the number of H-2A workers requested by employers in a county.[8]

3.4.3 Mechanisms

There are several potential explanations for the observed long-run effects of migration on labor markets. One possibility is the existence of potential complementarities between incumbent and newly-arrived workers that might lead to deeper specialization in tasks and higher productivity (Ottaviano and Peri, 2012; Peri and Sparber, 2009). Alternatively, agricultural firms can eventually

[8] This positive relationship between migration and the growth of the H-2A program could be indicative of seasonal guest worker programs leading to higher migration rates in the long-run. The public DOL data on individual H-2A requests made by employers does not record the place of origin of migrant workers being employed on each county under the program, so it is not possible to correlate the origin location of seasonal guest workers with the origin location of independent migration flows recorded in the MCAS data.

adjust their capital-labor ratios and technology choices, substituting production away from labor and focusing in less labor-intensive operations (Clemens et al., 2018). Another potential explanation might be that, due to the fact that the agricultural sector produces mainly tradable goods, increases in the relative abundance of labor might cause agricultural firms to make most of the adjustment through an expansion in output rather than through prices, as suggested by (Burstein et al., 2020).

We investigate which of these explanations are driving the observed long-run effects on wages and employment by analyzing how changes in migration rates differentially change agricultural production outcomes between 2007 and 2017 using data from the Census of Agriculture.

First, Table 3.7 presents the IV results of estimating equation (3.6) on outcomes related to the number of farms, farm operators, and crop choice. Columns 1, 2, and 3, show that counties with higher migration rates between 2007 and 2017 had as a result a higher total number farms, more total farm operators, and fewer total acres of cropland planted.[9] Farm operators are defined as individuals involved in the day-to-day decisions of the farm excluding all hired workers unless they were hired specifically as managers. The increase in the total number of operators and in operators per acre indicates a higher proportion of the agricultural workforce shifted towards more managerial roles in counties where migration flows were higher. This is consistent with a task-specialization process where increases in the supply of one type of worker allow other complementary types of workers to become more productive and earn higher wages.

Table 3.7: Long run impacts of migration – Task specialization and cultivation practices

			Area Harvested					Area Operated	
	(1) Number of farms	(2) Number of Operators	(3) Total cropland	(4) Field crops	(5) Fruits and tree nuts	(6) Horticulture	(7) Vegetables	(8) Family operated	(9) Non-family operated
Migration Rate (m_c)	1.794***	1.561**	-3.709***	-3.421***	1.650	7.631*	10.382**	-2.259	-18.555***
	(0.586)	(0.634)	(1.015)	(1.069)	(1.972)	(4.318)	(4.966)	(2.289)	(2.570)
Observations	3046	3045	3037	3011	1987	1002	1918	2849	2580
Kleinberg-Paap F	216.388	216.376	216.164	201.812	149.53	80.117	124.088	194.558	173.16

Notes: All outcomes are from the Census of Agriculture and are computed as the county-level log difference between 2007 and 2017. *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the county level in parentheses.

Regarding crop choice, columns 4 to 7 of Table 3.7 show that counties where migration rates fell most shifted away from labor-intensive crops (i.e. fruits and tree nuts, horticulture, and vegetables)

[9] The reduction in total agricultural area is also observed if the measure of total area operated (i.e. including pastures and woodland) is used instead of total cropland.

Table 3.8: Long run impacts of migration – Input substitution

	Machinery value - Owned				Machinery expenses - Rented			
	(1) Total Value	(2) $\frac{\text{Value}}{\text{acre}}$	(3) $\frac{\text{Value}}{\text{worker}}$	(4) $\frac{\text{Value}}{\text{farm}}$	(5) Total Value	(6) $\frac{\text{Value}}{\text{acre}}$	(7) $\frac{\text{Value}}{\text{worker}}$	(8) $\frac{\text{Value}}{\text{farm}}$
Migration Rate (m_c)	-1.812**	-0.682	-0.815	-3.533***	-5.496***	-4.983**	-3.875**	-7.496***
	(0.793)	(0.645)	(0.890)	(0.650)	(2.008)	(1.977)	(1.921)	(2.031)
Observations	3042	3028	2972	3042	2620	2615	2609	2620
Kleinberg-Paap F	260.401	248.476	249.431	260.401	202.052	195.672	201.614	202.052

Notes: All outcomes are from the Census of Agriculture and are computed as the county-level log difference between 2007 and 2017. *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the county level in parentheses.

Table 3.9: Long run impacts of migration – Productivity

	Farmland Value		Total Income Reported - All commodities			
	(1) Total Value	(2) $\frac{\text{Value}}{\text{acre}}$	(3) Total Value	(4) $\frac{\text{Value}}{\text{acre}}$	(5) $\frac{\text{Value}}{\text{worker}}$	(6) $\frac{\text{Value}}{\text{farm}}$
Migration Rate (m_c)	2.127**	3.378***	-3.054***	-2.595***	-2.039**	-5.073***
	(0.826)	(0.882)	(1.009)	(0.919)	(0.879)	(1.064)
N	3043	3030	2983	2972	2925	2983
Kleinberg-Paap F	216.096	206.269	205.693	196.106	197.014	205.693

Notes: All outcomes are from the Census of Agriculture and are computed as the county-level log difference between 2007 and 2017. *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the county level in parentheses.

and experience instead an increase in total acreage devoted to field crops which, in general, demand fewer workers per acre. Consistent with this change in crop composition patterns, results shown in table 3.8 show that counties where exogenous increases in migration rates took place had as a consequence lower rates of mechanization, measured both as the reported asset value of all machinery owned by producers, and as the cost of machinery rentals. The impact of migration on mechanization remains negative whether measured as total value, or in proportion to total acres harvested, workers hired, or farms in operation. This result is consistent with other evidence that show how reductions in the supply of low-wage labor lead to long-run mechanization (Hornbeck and Naidu, 2014).

Finally, the results presented in Table 3.9 explore if the positive effect of migration on wages can be explained through increases in land productivity or total ouptut. Columns 1 and 2 reveal that higher migration rates cause average farmland values in a county, measured in dollars per acre, to increase in the long run. This increase in farmland values is indicative of higher land productivity and suggests that increasing migration rates, perhaps by expanding the pool of potential employees allows producers to select more skilled workers that lead to productivity increases. However, columns 3 to 6 of the same table show that higher migration rates caused reductions in the average income reported by farmers, irrespective of whether it is measured in proportion to land cultivated, number of workers hired, or number of agricultural operations.[10] These results suggest that increased migration rates do not seem to translate into expansions of agricultural output.

3.5 Conclusion

Industry-wide labor shortages and migratory policy are two extensively discussed economic issues in current policy debates. This paper contributes to that debate by quantifying the relative importance of different margins of adjustment through which employers are responding to changes in labor supply, and by showing how these adjustments change over time. Our results quantify the effect that changes in migration flows from Mexico have on the organization of the U.S. agricultural sector. We show the specific way in which agricultural labor markets and farmers' production decisions respond to fluctuations in migrant labor supply, and uncover stark differences in these responses between the short and the long run. While yearly reductions in migration rates push up agricultural wages and lead producers to compensate labor scarcity by increasing their demand for seasonal guest workers, in the long run local economies with fewer migrant arrivals experience broader changes in their agricultural industry. These changes—related to producers crop choices and production practices—have led counties with secular slowdowns in migration rates to experience decreases in agricultural employment and agricultural salaries, as well as reductions in average farmland values. Understanding if these migration-driven changes in production have economically important effects on the price of agricultural products faced by consumers, and whether changes in agricultural

[10]Similar results are obtained if total revenues (i.e. total value of sales of agricultural products) is used instead of income.

productivity might have spillover effects on other sectors of the economy where a large share of workers are also foreign immigrants are two potentially fruitful avenues for future research.

Bibliography

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

Brown, R. and Velásquez, A. (2017). The effect of violent crime on the human capital accumulation of young adults. *Journal of development economics*, 127:1–12.

Burstein, A., Hanson, G., Tian, L., and Vogel, J. (2020). Tradability and the labor-market impact of immigration: Theory and evidence from the united states. *Econometrica*, 88(3):1071–1112.

Caballero, M. E., Cadena, B. C., and Kovak, B. K. (2018). Measuring geographic migration patterns using matrìculas consulares. *Demography*, 55(3):1119–1145.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Castillo, J. C., Mejía, D., and Restrepo, P. (2020). Scarcity without leviathan: The violent effects of cocaine supply shortages in the mexican drug war. *Review of Economics and Statistics*, 102(2):269–286.

Charlton, D. and Castillo, M. (2022). Housing booms and h-2a agricultural guest worker employment. *American Journal of Agricultural Economics*, Forthcoming.

Charlton, D. and Taylor, J. E. (2016). A declining farm workforce: Analysis of panel data from rural mexico. *American Journal of Agricultural Economics*, 98(4):1158–1180.

Clemens, M. A. (2021). Violence, development, and migration waves: Evidence from central american child migrant apprehensions. *Journal of Urban Economics*, 124:103355.

Clemens, M. A., Lewis, E. G., and Postel, H. M. (2018). Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion. *American Economic Review*, 108(6):1468–87.

Dell, M. (2015). Trafficking networks and the mexican drug war. *American Economic Review*, 105(6):1738–79.

Gonzalez-Barrera, A. (2015). More mexicans leaving than coming to the u.s. *Pew Research Center*, 19.

Guerrero-Gutiérrez, E. (2011). Security, drugs, and violence in mexico: A survey. In *7th North American Forum, Washington DC*. sn.

Hanson, G. H. and McIntosh, C. (2010). The Great Mexican Emigration. *The Review of Economics and Statistics*, 92(4):798–810.

Hernandez, T. and Gabbard, S. (2019). Findings from the national agricultural workers survey (naws) 2015–2016. a demographic and employment profile of united states farmworkers. *Department of Labor, Employment and Training Administration, Washington; 2019. Report No. 13*.

Hornbeck, R. and Naidu, S. (2014). When the levee breaks: Black migration and economic development in the american south. *American Economic Review*, 104(3):963–90.

Imbert, C., Seror, M., Zhang, Y., and Zylberberg, Y. (2022). Migrants and firms: Evidence from china. *American Economic Review*, 112(6):1885–1914.

Kleemans, M. and Magruder, J. (2018). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. *The Economic Journal*, 128(613):2032–2065.

Martin, P. L. (2020). Immigration and farm labor: Challenges and opportunities. In Martin, P., Goodhue, R., and Brian, W., editors, *California agriculture: dimensions and issues*. Giannini Foundation of Agricultural Economics.

Massey, D. S., Rugh, J. S., and Pren, K. A. (2010). The geography of undocumented mexican migration. *Mexican studies/Estudios mexicanos*, 26(1):129–152.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics*, 118(2):549–599.

Munshi, K. (2014). Community networks and migration. In *The Oxford handbook of the economics of networks*.

Orozco-Aleman, S. and Gonzalez-Lozano, H. (2018). Drug violence and migration flows lessons from the mexican drug war. *Journal of Human Resources*, 53(3):717–749.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, 10(1):152–197.

Passel, J. S. and Cohn, D. (2018). Us unauthorized immigrant total dips to lowest level in a decade. *Pew Research Center*, 17.

Peri, G. and Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3):135–69.

Peri, G. and Zaiour, R. (2022). Labor shortages and the immigration shortfall. *Econofact*.

Rutledge, Z. and Mérel, P. (2023). Farm labor supply and fruit and vegetable production. *American Journal of Agricultural Economics*, 105(2):644–673.

Tian, Y., Caballero, M. E., and Kovak, B. K. (2022). Social learning along international migrant networks. *Journal of Economic Behavior & Organization*, 195:103–121.

Tiburcio, E. and Camarena, K. R. (2023). The local reaction to unauthorized mexican migration to the us. *Working Paper*.

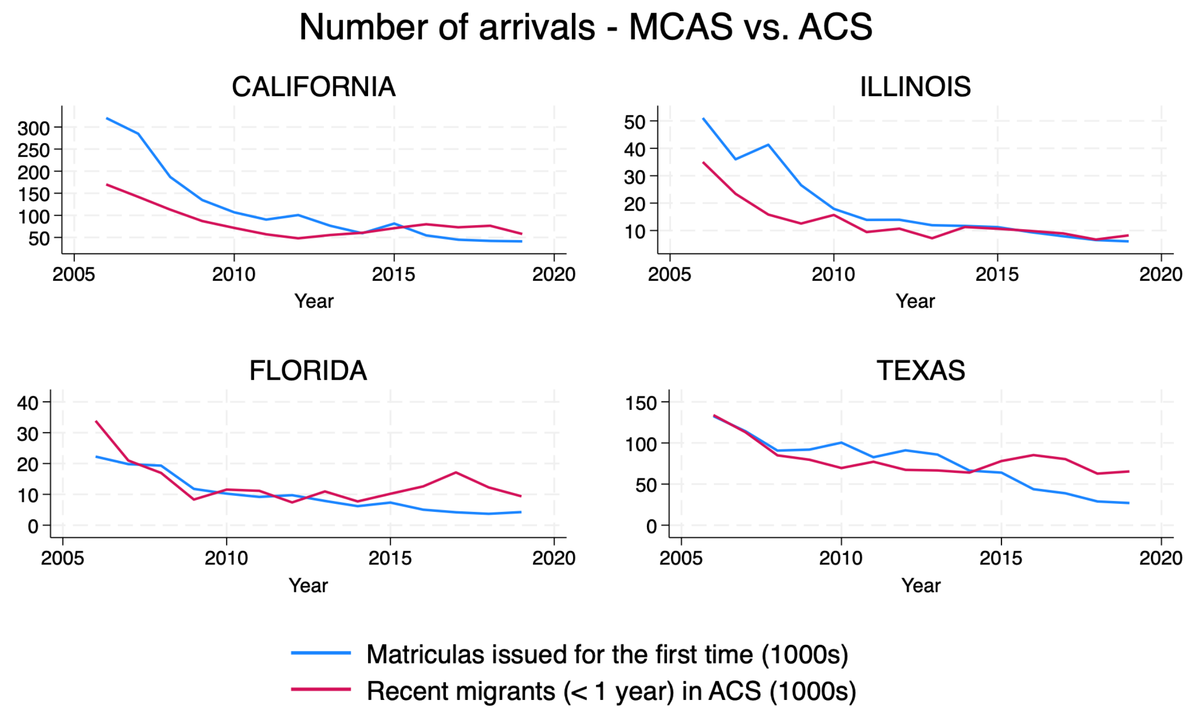
Velásquez, A. (2020). The economic burden of crime evidence from mexico. *Journal of Human Resources*, 55(4):1287–1318.

Williams, P. (2012). The terrorism debate over mexican drug trafficking violence. *Terrorism and Political Violence*, 24(2):259–278.

Zahniser, S., Taylor, J. E., Hertz, T., and Charlton, D. (2018). Farm labor markets in the united states and mexico pose challenges for us agriculture. *United States Department of Agriculture, Economic Research Service.*

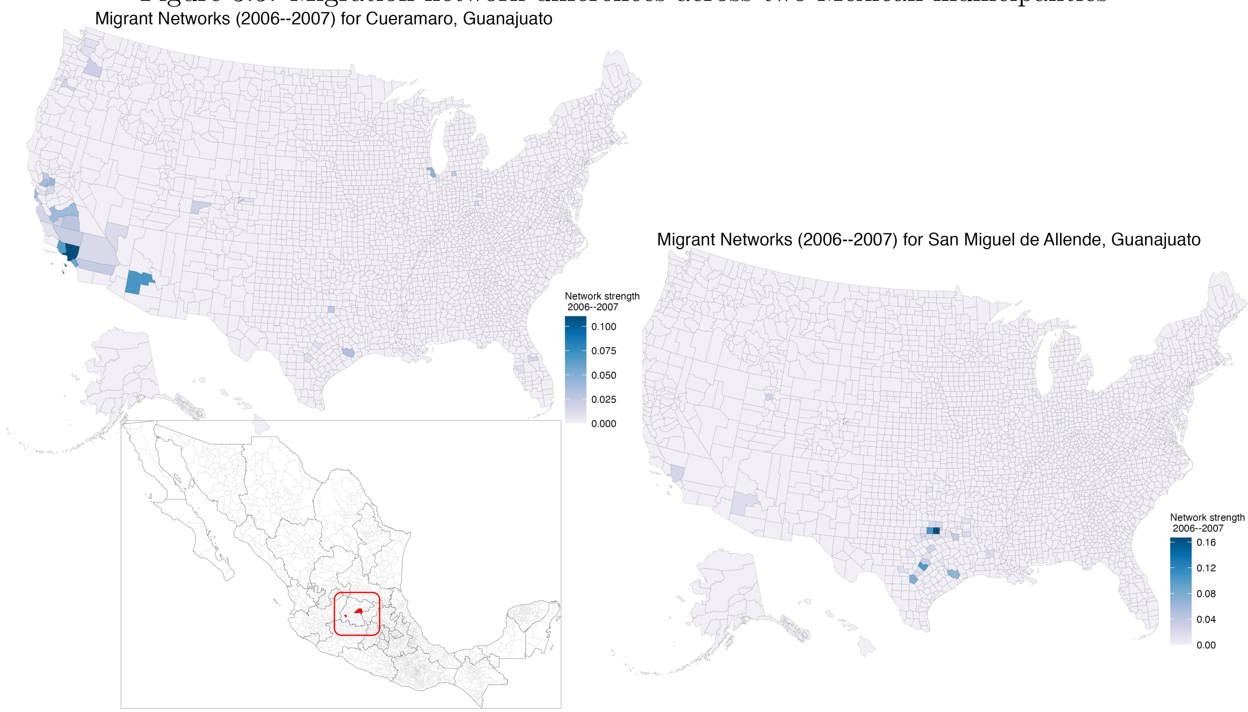
3.6 Appendix A: Additional tables and figures

Figure 3.4: Migration inflows across data sources – State level trends



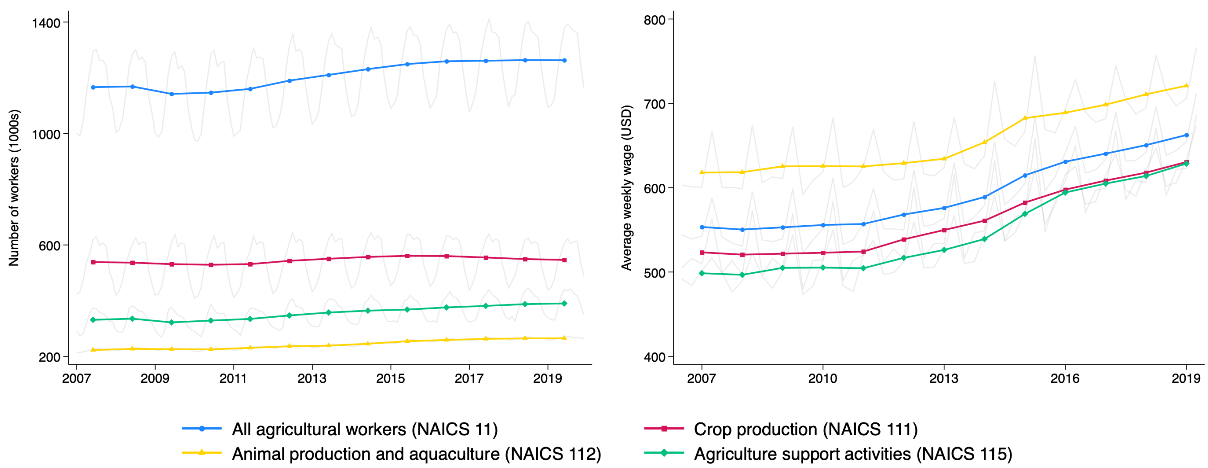
Notes: Number of yearly arrivals from Mexico comparing MCAS and ACS. Blue lines show the number of Matriculas Consulares de Alta Seguridad (*MCAS*) issued for the first time by all Mexican consulates in each state. Red lines show the number of migrants recorded in the ACS as having moved from Mexico to each state within the previous year.

Figure 3.5: Migration network differences across two Mexican municipalities



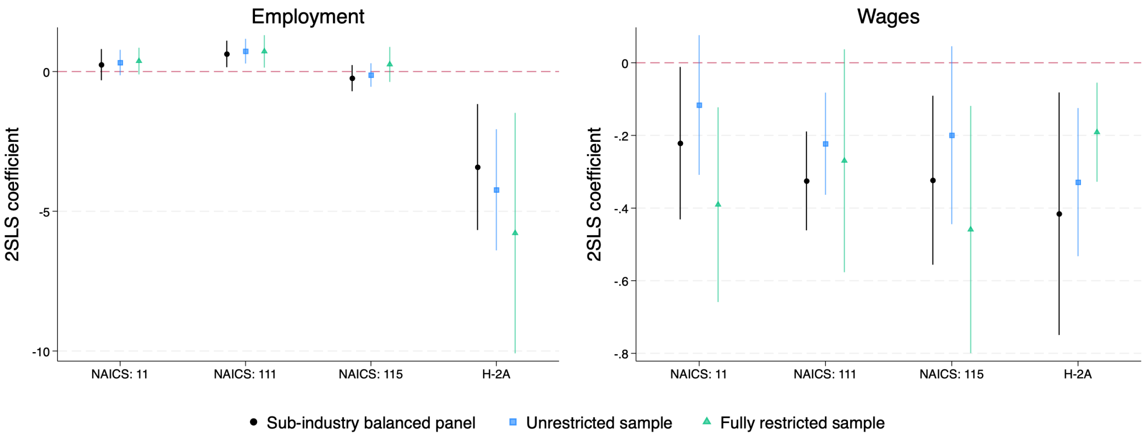
Notes: County-municipality migration network strength measured as the ratio between the number of migrants originated in each Mexican municipality moving to each U.S. county over the total number of migrants in the municipality. Left panel: Migration network for the municipality of Cueramaro, Guanajuato. Right panel: Network for San Miguel de Allende, Guanajuato.

Figure 3.6: Agricultural labor – National level trends

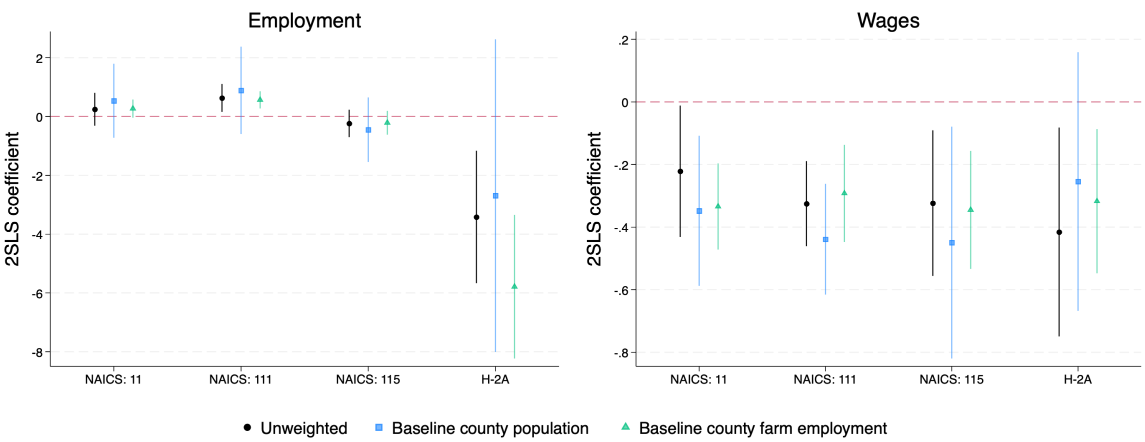


Notes: Data from the QCEW. Gray lines represent quarterly figures; solid color lines yearly averages.

Figure 3.7: Short run results – Sensitivity to alternative estimation samples and weights



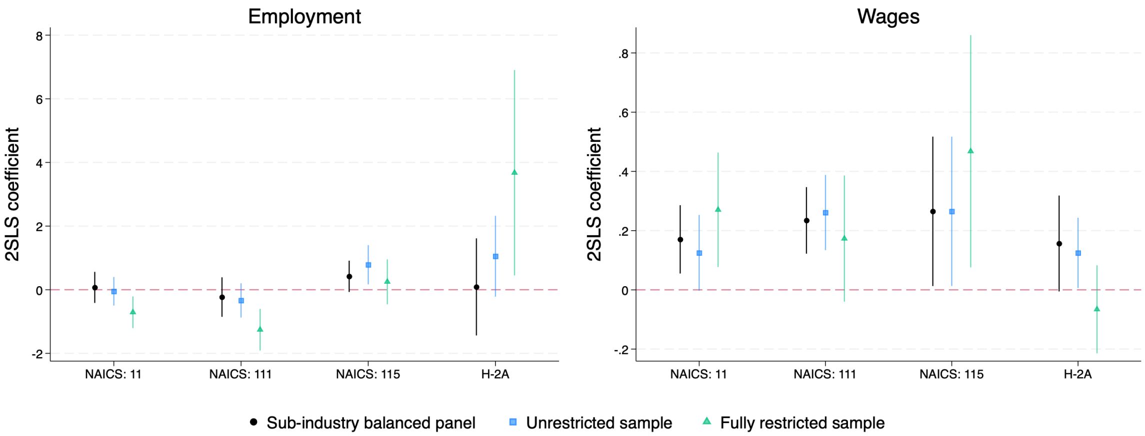
(a) Alternative estimation samples



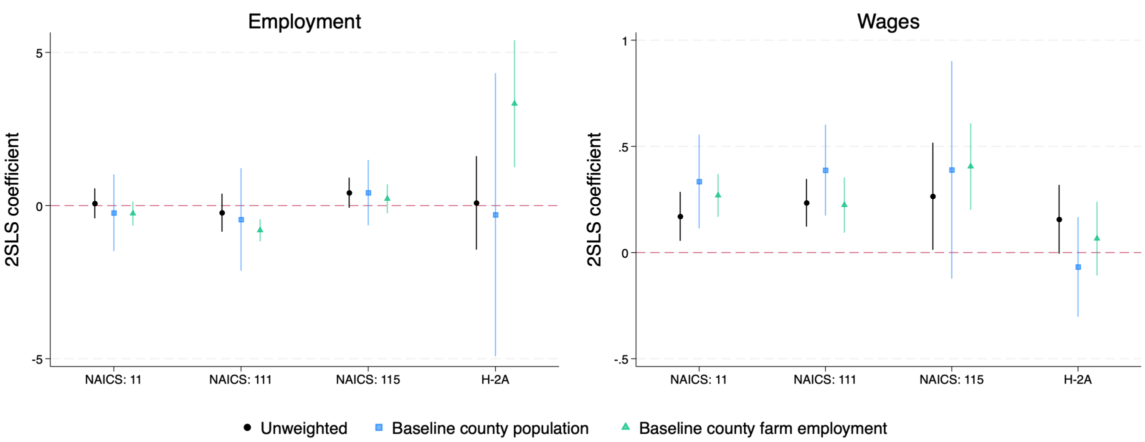
(b) Alternative regression weights

Notes:

Figure 3.8: Long run results – Sensitivity to alternative estimation samples and weights



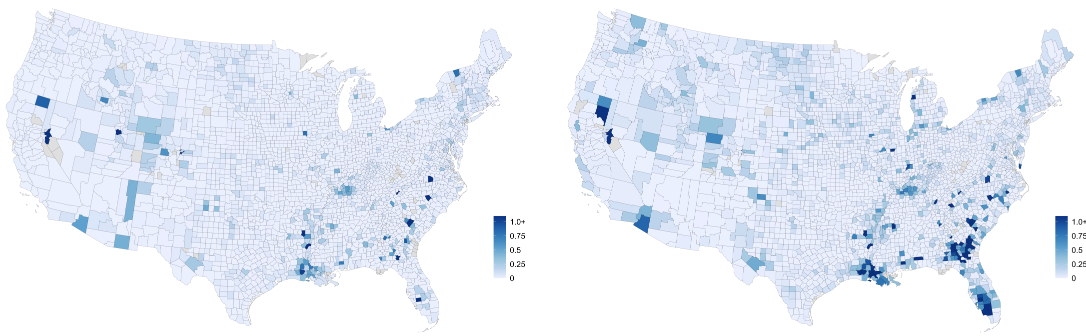
(a) Alternative estimation samples



(b) Alternative regression weights

Notes:

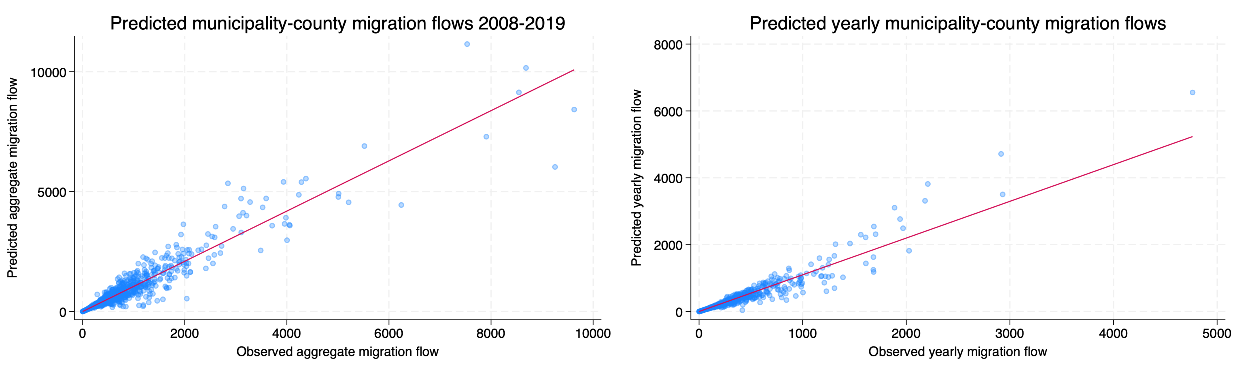
Figure 3.9: H-2A Visa requests by county – 2008-2018



(a) 2008 (b) 2018

Notes: Number of H-2A requests as fraction of total farm employment in county based on the QCEW. Individual H-2A request data from the Department of Labor (*DOL*).

Figure 3.10: Observed vs. network-predicted migration flows



Notes: Left panel: binscatter of aggregate (2008-2019) municipality-county observed migration flows and the predicted flow obtained from multiplying total municipality outflows by the network strength measure $\phi_{o,c}$. Right panel: yearly predicted and observed municipality-county migration flows.

3.7 Appendix B: Additional First-stage results

To test if the negative relationship between the instrument and migration rates is driven by the 'share' component of the instrument, we define the alternative instrumental variable

$$Z^I_{c,t} = \frac{1}{P_{c,t^0}} \sum_m \left[\text{Homicides}_{m,t} \times \mathbb{1}(\phi^{t^0}_{m,c} > 0) \right]$$

where all (non-zero) origin-destination links are weighted equally. The results of re-estimating regression 3.4 using $Z^I_{c,t}$ as an instrument are displayed in Table 3.10 and show that the negative sign is still present once county fixed effects are included.

Table 3.10: First-stage estimates – Instrumental variable with no network component

	Yearly immigration rate ($m_{c,t}$)			
	(1)	(2)	(3)	(4)
$Z^I_{c,t}$: No migrant network IV	0.004***	0.004***	-0.003***	-0.002***
	(0.000)	(0.000)	(0.001)	(0.000)
Observations	37680	37680	37680	37680
Counties	3140	3140	3140	3140
Year FE	No	Yes	No	Yes
County FE	No	No	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the municipality level in parentheses.

It might also be that the negative relationship observed arises because the year-to-year migration measure is too noisy at such a fine temporal disaggregation. This could happen if for a large number of migrants the arrival date into the U.S. and the decision to get a MCAS card are years apart. While the aggregate trends shown in Figures 3.1 and 3.4 do not seem to suggest this, we nonetheless test this by running an alternative version of regression 3.4, where both migration rates and the instrument are aggregated into three-year bins. Once again, results for this exercise —shown in Table 3.11— still exhibit the change in sign once county fixed effects are included.

Table 3.11: First-stage estimates – Three-year migration rate aggregation

	3-Year immigration rate			
	(1)	(2)	(3)	(4)
$Z^{3-year}_{c,t}$: Violence shift-share IV	8.039***	8.252***	-7.871***	-6.489***
	(0.473)	(0.494)	(1.143)	(1.051)
Observations	12568	12568	12568	12568
Counties	3142	3142	3142	3142
Year FE	No	Yes	No	Yes
County FE	No	No	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the municipality level in parentheses.

We finally show that the inclusion of unit fixed effects renders the relationship between violence and migration inflows negative even when measured at the origin-destination pair level. Given that the MCAS data allows us to observe the magnitude of all migration flows originating in every Mexican municipality headed to each U.S. county, we are able to run the following regression

$$\frac{M_{c,m,t}}{Pop_{m,t^0}} = \beta_0 + \beta_1 \left[\text{Homicides}_{m,t} \times \phi^{t^0}_{m,c}\right] + \delta_t + \gamma_c + \eta_m + \chi_{c,m} + \varepsilon_{c,m,t} \tag{3.7}$$

where $M_{c,m,t}$ is the observed migration from c to m at t, and γ_c, η_m, and $\chi_{c,m}$ are, respectively, county, municipality, and county-by-municipality fixed effects.

Results for regression equation 3.7 are shown in Table 3.12. These results show that while the separate inclusion of either municipality or county fixed effects does not affect the cross-sectional positive relationship between violence and migration, once municipality-by-county fixed effects are included the relationship once again changes sign and becomes negative.

Table 3.12: First-stage estimates – Origin-destination level regressions

	Yearly Origin-destination immigration rate ($m_{o,m,t}$)					
	(1)	(2)	(3)	(4)	(5)	(6)
Homicides$_{o,t} \times \phi_{o,c}$	0.645***	0.655***	0.731***	0.226**	0.310**	-0.126*
	(0.238)	(0.243)	(0.267)	(0.111)	(0.136)	(0.067)
Observations	5099796	5099796	5099796	5099796	5099796	5099796
Year FE	No	Yes	Yes	Yes	Yes	Yes
Municipality FE	No	No	Yes	No	Yes	No
County FE	No	No	No	Yes	Yes	No
Muni × County FE	No	No	No	No	No	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.10. Standard errors clustered at the municipality level in parentheses.

3.8 Appendix C: Violence and Migration - Simulation Exercise

This section describes a simulation exercise that shows how the change in sign of the first-stage regression coefficients shown in Section 3.3 is consistent with a data generating process where violence-driven origin-destination migration flows are determined by the combination of long-run violence shocks that tend to follow the migrant networks and short-run shocks that follow an independent decision rule.

Setup:

Let the world consist of I origin municipalities indexed by i, J destination counties indexed by j, and T years indexed by t. Each municipality is affected violence that induces outmigration, and migrants select a destination according to some decision rule.

Define:

V_{it}: Violence in municipality i in year t

M_{ijt}: Migration from municipality i to county j in year t.

N_{ij}: Historical migration network. Normalized to sum to 1 within i: $\sum_j N_{ij} = 1, \forall i$.

D_{ij}: Alternate decision rule for destination selection. Also normalized to $\sum_j D_{ij} = 1, \forall i$.

Additionally, for any variable presented without a subscript, let it represent the sum over that subscript. E.g. $V_i \equiv \sum_t V_{it}$.

Yearly violence in a municipality is modeled as a long-run average plus a short-run shock:

$$V_{it} = \bar{V}_i + \tilde{V}_{it}$$

where

$$\bar{V}_i \overset{i.i.d}{\sim} \mathcal{N}(0, Var(\bar{V}_i))$$
$$\tilde{V}_{it} \overset{i.i.d}{\sim} \mathcal{N}(0, Var(\tilde{V}_{it})).$$

Assume outmigration from i to j induced by long-run violence follows the network decision rule N_{ij}, while outmigration caused by short-run violence follows some independent short-run decision rule D_{ij}. That is, migration flows are determined by the following data generating process:

$$M_{ijt} = \beta^L N_{ij} \bar{V}_i + \beta^S D_{ij} \tilde{V}_{it} + \alpha t + \epsilon_{ijt} \tag{3.8}$$

For some normally-distributed mean-zero i.i.d ϵ, and for variables N_{ij}, D_{ij} distributed such that:

$$\begin{pmatrix} log(N_{ij}) \\ log(D_{ij}) \end{pmatrix} \overset{i.i.d}{\sim} \mathcal{N} \begin{pmatrix} 1 & \sigma^2_{dn} \\ \sigma^2_{dn} & 1 \end{pmatrix}$$

where σ^2_{dn} determines the correlation between long-run and short-run decision rules.

Note that, based on (3.8),

$$M_{it} = \sum_j M_{ijt} = \beta^L \bar{V}_i + \beta^S \tilde{V}_{it} + \sum_j \epsilon_{ijt},$$

since $\sum_j N_{ij} = \sum_j D_{ij} = 1$.

The within-i transformation of M_{it} yields:

$$\dot{M}_{it} \equiv M_{it} - \frac{1}{T}\sum_t M_{it} = \beta^S \left[\tilde{V}_{it} - \frac{1}{T}\sum_t \tilde{V}_{it}\right] + \left(\sum_j \epsilon_{ijt} - \sum_t \sum_j \epsilon_{ijt}\right),$$

and so the parameter β^S can thus be recovered through a regression of $\dot{M}_{it}$ on V_{it} or, equivalently, through a municipality fixed effect of the form:

$$\beta^S: \quad M_{it} = \beta V_{it} + \gamma_i + \delta_t + \epsilon_{it} \tag{3.9}$$

Similarly, aggregating across all time periods yields:

$$\begin{aligned} M_i &= \sum_t M_{it} = T\beta^L \bar{V}_i + \beta^S \sum_t \tilde{V}_{it} + \sum_t \sum_j \epsilon_{ijt} \\ V_i &= \sum_t V_{it} = \sum_t [\bar{V}_i + \tilde{V}_{it}] = T\bar{V}_i + \sum_t \tilde{V}_{it}, \end{aligned}$$

and thus the parameter β^L can be recovered by a regression of the form:

$$\beta^L: \quad M_i = \beta V_i + \epsilon_i, \tag{3.10}$$

since $E[\tilde{V}_{it}] = 0$, and thus $V_i/T \to \bar{V}_i$ as $T \to \infty$.

Calibration:

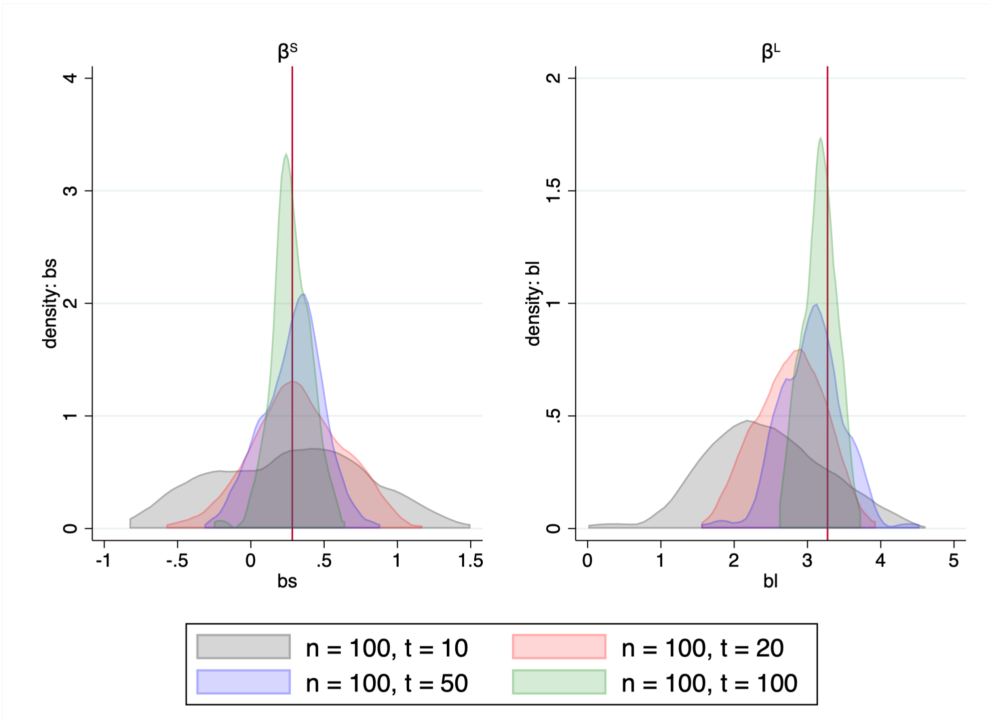
Using the expressions above, we calibrate β^S, β^L, and the variance of the violence shocks from the data:

$$\begin{aligned} \beta^S &= 0.283 \\ \beta^L &= 3.274 \\ Var(\bar{V}_i) &= 0.00000007317037 \\ Var(\tilde{V}_{it}) &= 0.00000014502362 \end{aligned}$$

Simulation:

With the calibrated parameters, we simulate data for an equal number of equally-sized municipalities and counties. Figure 3.11 shows the resulting distribution of estimating β^S and β^L following equations (3.10) and (3.9) in 100 different data draws and at varying levels of T. The figure confirms that regression (3.10) recovers β_L as T grows.

Figure 3.11: Distribution of β^L and β^S for 100 regressions - Fixed number of municipalities



As a check, we estimate the out-migration regressions (3.9) and (3.10) For a single data draw and confirm the estimates are close to those observed in true data. Results for this exercise are shown in Tables 3.13 and 3.14.

County-level migration inflows:

Following our empirical strategy, we also aggregate the municipality level simulated outflows to the county level:

$$Z_{jt} \equiv \sum_i (N_{ij} \times V_{it})$$

And, estimate a regression of M_{jt} on Z_{jt}:

Table 3.13: True data: Homicides and migration yearly correlation at the Mexican municipality level

	(1) m_{it}	(2) m_{it}	(3) m_{it}	(4) m_{it}	(5) m_i
Yearly Homicide Rate	0.956***	1.293***	-0.213	0.283**	
	(0.176)	(0.188)	(0.178)	(0.138)	
Aggregate Homicide Rate					3.274***
					(0.457)
Observations	29232	29232	29232	29232	2436
R^2	0.004	0.154	0.000	0.359	0.029
Year FE	No	Yes	No	Yes	No
Municipality FE	No	No	Yes	Yes	No

*** p<0.01, ** p<0.05, * p<0.10

Table 3.14: Simulated data: Violence and outflows

	(1) M_{it}	(2) M_{it}	(3) M_{it}	(4) M_{it}	(5) M_i
V_{it}	0.693***	0.794***	0.172	0.292***	
	(0.195)	(0.083)	(0.211)	(0.061)	
V_i					3.004***
					(0.147)
Observations	10000	10000	10000	10000	100
R^2	0.001	0.892	0.000	0.895	0.813
Year FE	No	Yes	No	Yes	No
Municipality FE	No	No	Yes	Yes	No

*** p<0.01, ** p<0.05, * p<0.10. $I = 100$; $J = 100$; $T = 100$

$$\begin{aligned} M_{jt} &= \beta \underbrace{\sum_i (N_{ij} \times V_{it})}_{\equiv Z_{jt}} + \gamma_j + \varepsilon_{jt} \\ &= \beta \sum_i N_{ij} \times (\bar{V}_i + \tilde{V}_{it}) + \gamma_j + \varepsilon_{jt} \\ &= \beta \sum_i N_{ij} \bar{V}_i + \beta \sum_i N_{ij} \tilde{V}_{it} + \gamma_j + \varepsilon_{jt} \end{aligned} \tag{3.11}$$

Results from estimating regression equation (3.11) on a single draw of simulated data with $\sigma^2_{dn} = 0$ are reported in Table 3.15. Consistent with the estimations made on real data, the results of this simulation exercise show that even if the correlation between municipality-level violence and outmigration is *defined* to be positive, a negative correlation between the county-aggregated violence variable Z_{jt} and migration inflows can arise. We take this to mean that the underlying relationship between migration and violence driving our results is consistent with a data generating process where violence-driven origin-destination migration flows are determined by the combination of long-run violence shocks that tend to follow the migrant networks and short-run shocks that follow an independent decision rule.

Table 3.15: Simulated data: Violence and inflows: $Corr(N_{ij}, D_{ij}) = 0$

	(1) M_{jt}	(2) M_{jt}	(3) M_{jt}	(4) M_{jt}	(5) M_j
Z_{jt}	-4.099***	-0.194	-5.122***	-1.011*	
	(0.922)	(0.460)	(1.040)	(0.517)	
Z_j					3.810***
					(1.143)
Observations	10000	10000	10000	10000	100
R^2	0.001	0.895	0.002	0.896	0.099
Year FE	No	Yes	No	Yes	No
Municipality FE	No	No	Yes	Yes	No

*** p<0.01, ** p<0.05, * p<0.10. $I = 100$; $J = 100$; $T = 100$; $\sigma^2_{dn} = 0$.