Farm Size Distribution, Weather Shocks, and Agricultural Productivity

Julián Arteaga
World Bank

Nicolás de Roux
Universidad de Los Andes

Margarita Gáfaro
Banco de la Republica

Ana María Ibáñez
IADB

Heitor S. Pellegrina
University of Notre Dame *

January 6, 2025

Abstract

This paper studies the dynamics of farm size distribution, how they are influenced by weather shocks, and the implications for aggregate productivity. Using data from several developing countries, we first document new empirical facts about households' landholding choices and how weather shocks influence these decisions. Building on a rich longitudinal dataset for Colombia on farm sizes, land transactions, and households' consumption and investment decisions, we then show that weather shocks increase the frequency of land sales and reduce farm sizes within municipalities, especially among smaller farms. To rationalize these facts, we develop a dynamic, heterogeneous household model in which uninsured farmers make landholding and occupational choices. Our calibrated model shows that uninsured risk substantially curbs aggregate agricultural productivity, and that the effects of temporary weather shocks on farm size and agricultural output are highly persistent, taking more than a decade to fade out.

*We thank Paola Poveda, Juliana Quigua, Salvador Traettino, and Xinyi Li for providing excellent research assistance. We also thank participants in presentations at Barcelona Summer Forum, Columbia, University of Chicago, RIDGE, NEUDC, EESP/FGV, Puc-Rio, LACEA, PacDev, Northeastern, World Bank, University of Chile, INSPER, UC-Davis, AAEA, IEA World Congress, University of Notre Dame, and Universidad de Los Andes. The findings, interpretations, and conclusions expressed in this paper are solely those of the authors and do not necessarily represent the views of the World Bank, IADB or Banco de la Republica. Email: jarteaga@worldbank.org, nicolas.de.roux@uniandes.edu.co, mgafargo@banrep.gov.co, anaib@iadb.org, hpelleg3@nd.edu

1 Introduction

In low- and middle-income countries, a large share of households live in small, low-productivity farms (Restuccia et al., 2008; Gollin et al., 2014; Herrendorf et al., 2014). Moreover, the lack of insurance mechanisms in these countries leaves rural households largely exposed to the weather shocks inherent to agricultural production, strongly influencing their consumption and investment decisions. Prior research has shown that the farm size distribution in these countries—which is characterized by a large fraction of small farms—significantly constrains aggregate productivity (Adamopoulos and Restuccia, 2014; Foster and Rosenzweig, 2022). We lack, however, better understanding of how the dynamics of households consumption and investment decisions contribute to the predominance of small farms, and the influence of weather shocks on these decisions.

In this paper, we study the dynamics of farm size distribution, how they are influenced by weather shocks, and the implications for aggregate agricultural productivity. We first document new empirical facts about households landholding decisions and their response to weather shocks across developing countries, by bringing together household-level data from Colombia and eight West African countries. In addition, building on two unique administrative panel datasets from Colombia—one containing hundreds of thousands of land transactions across the country, and another with the universe of land-holding registrations—we estimate how land-market transactions and the farm size distribution are influenced by weather shocks at the municipality level. To rationalize the empirical facts we uncover, we develop and calibrate a dynamic, heterogeneous household model, which we use to study quantitatively the dynamic effects of weather shocks and implications for aggregate productivity.

Our paper begins by documenting two empirical facts on the relationships between farm size, farm size growth, and exit rates that resemble empirical relationships found in the firm size distribution literature based on the manufacturing sector—e.g., Jaimovich et al. (2023) and Clementi and Palazzo (2016). First, there is more churning among households with smaller farms: i.e., small farms are more likely to exit farming, but surviving small farms have higher farm size growth when compared to surviving large ones. Second, when farmers experience a weather shock, there is an increase in the exit probability across farms of all sizes. Among surviving farms, smaller farms tend to shrink when they experience the weather shock, while large farms tend to grow.

To better understand the implications of these household-level decisions on the farm size distribution, we turn to our data from Colombia, which give us the opportunity of observing the universe of farms as well as detailed information on the history of land transactions.

Using these data, we first show that weather shocks substantially increase the number land transactions in a municipality: specifically, they increase the number of plot sales (full or partial) and the number or mortgage contracts. Results also indicate that this increase in plot sales is largely driven by purchases made by landless households. Second, we find that weather shocks increase the total number of farmers in a municipality, and that this result comes mostly from an increase in the total number of smaller farms—in other words, weather shocks lead to land fragmentation. Third, we lastly show that, when experiencing a weather shock, households reduce their consumption, increase their probability of migrating to another municipality, and reduce their ownership of durable assets other than land. Importantly, throughout our analysis, we provide a host of robustness tests, in which we experiment with alternative sets of controls, measures of weather shocks, definitions of farm size, and controls capturing alternative mechanisms.

Motivated by these empirical facts, we develop a dynamic, heterogeneous household model. In our framework, households are either farmers or workers and, in each period, they make decisions about their occupation in the subsequent period. If they choose to farm they must decide how much land to acquire for production in advance of the harvest season. Households are heterogeneous in terms of their landholdings, and in terms of their farming and non-farm working productivity. These productivities are subject to shocks, which generates uncertainty. Based on the reality of rural markets in developing countries, households lack access to both agricultural insurance and land rental markets. In this context, landholdings play a dual role: they serve as a form of wealth accumulation and also as a tool for consumption smoothing. The landholding decisions, together with the entry and exit decisions, jointly determine the farm size distribution in every period.

In the model's steady state, a larger proportion of small farms exit farming compared to large farms, as the outside option of becoming a worker is relatively more attractive to smaller farms. To investigate household behavior in response to weather shocks, we analyze the transition dynamics. Specifically, we simulate an unexpected negative aggregate productivity shock affecting only farmers. If the shock is sufficiently temporary—i.e., if the shocks has low persistence—, then the number of farmers in the economy rises, leading to a reduction in average farm size, replicating what we find in our empirical facts. This occurs because the negative shock induces some farmers, particularly small ones, to exit, as they place a higher value on smooth consumption. This leads to a reduction in the price of land. However, when the shock is temporary, the value of becoming a farmer remains unaffected for landless households, since future productivity remains fixed. Quantitatively, the entry of landless households can exceed the exit of small farmers. In contrast, when we simulate shocks with greater persistence, fewer landless households choose to enter farming, which tends to reduce the number of farmers in the economy, therefore increasing the average farm size.

To quantify the aggregate implications of the farm size dynamics to agricultural productivity, we calibrate the steady state of our model to match several statistics about the farm size distribution in Colombia. Our model matches the average and the variance of farm size, the share of households in agriculture, and the average farm size of new entrants. Our quantitative model indicates the existence of substantial misallocation in steady state, in which less productive farms have low marginal product of land—i.e., they own more land than optimal. Relative to the optimal allocation of land across farmers with heterogeneous farming productivity, total agricultural production is 36 percent lower.

We close our paper by studying the transition dynamics generated by weather shocks. These shocks generate significant impacts on farm size dynamics, consistent with our findings from Colombia. A 25% reduction in aggregate agricultural productivity results in a 20% drop in land prices, a 25% decline in agricultural output, and an influx of new, less productive landowners. The shock induces land fragmentation, reducing the average farm size by 1.2% and causing a 0.7% decrease in the average skill of farmers. Notably, the effects are highly persistent: while aggregate agricultural productivity recovers within five years, approximately 30 percent of the reductions in average farm size and farmer skill remain 20 years after the shock. These dynamics underscore the dual role of land as both a productive asset and a consumption-smoothing mechanism, as well as the far-reaching implications of weather shocks for the farm size distribution and agricultural productivity.

Related Literature. This paper contributes to a growing literature on the causes and consequences of the farm size distribution in developing countries. Previous studies have explored various driving factors, including the role of land institutions (Bolhuis et al., 2021; Adamopoulos and Restuccia, 2020, 2014; Adamopoulos et al., 2022; Chen et al., 2022), land- and labor-market imperfections (Foster and Rosenzweig, 2022; Acampora et al., 2022), and access to urban centers (Gáfaro and Pellegrina, 2022; Pellegrina, 2022; Rao et al., 2022; Madhok et al., 2022). Our contribution here is twofold. First, we document new facts about the dynamics of the farm size distribution—drawing a parallel to the literature on firm size distribution based on manufacturing—and about the response of the farm size distribution to weather shocks.[1] Second, we bring in a dynamic perspective to this literature, by analyzing how households consumption and investment decisions shape the farm size distribution in the absence of insurance and land markets.

[1] 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). To the best of our knowledge, we are the first to show the impact on the farm size distribution.

Importantly, our results 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). Notably, our results indicate that short-term shocks can have especially persistent effects on the economy.

By developing a heterogeneous household model for the agricultural sector, our paper speaks to the rich literature on heterogeneous firm models that builds on the seminal work of Hopenhayn (1992) and Aiyagari (1994). In particular, our work complements recent papers applying this class of models to agriculture. Brooks and Donovan (2020) study the implications of risk reduction induced by the construction of bridges. Manysheva (2022) and Gottlieb and Grobovšek (2019) focus on the role of communal land systems on aggregate agricultural productivity. Peralta-Alva et al. (2023) quantify the costs of tax revenues in low-income countries with large agricultural sectors. Mazur and Tetenyi (2024) investigate the macro-economic implications of agricultural input subsidies. Our key contribution here is to treat the farm size distribution as endogenous, determined by households consumption and investment decisions. Our approach to landholding choices is similar to papers modelling firm investments under adjustment costs—e.g., Khan and Thomas (2008). Moreover, with exception of Brooks and Donovan (2020), these recent applications focus on the stationary equilibrium, whereas here the transition dynamics is a core object of our analysis. Our work benefits from recent methodological advances, specifically the Sequence Space Jacobian method developed in Auclert et al. (2021), to solve for the transition dynamics.

Lastly, this paper contributes to the literature exploring the effects of weather shocks on agriculture. Negative productivity shocks often force poor landowners to deplete their assets to smooth consumption (Rosenzweig and Wolpin, 1993; Carter and Zimmerman, 2003; Kazianga and Udry, 2006). Farmers' responses to weather shocks can also include adjustments in labor and intermediate inputs use, loan repayment, changes in crop choice, migration, or investment in human capital (Jayachandran, 2006; Jessoe et al., 2018; Colmer, 2021; de Roux, 2021; Jagnani et al., 2021; Aragón et al., 2021).[2] 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

[2] In line with our findings, recent work by Kaboski et al. (2022) find that farmers invest in land when they win a large lottery.

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 Farm Size Dynamics and Weather Shocks across Countries

This section provides a first look at farm size dynamics and the impact of weather shocks across developing countries. To do so, we bring in household-level panel data for eight west African countries from the EHCVM surveys, plus the Colombian rural household panel data from ELCA, provided by Universidad de Los Andes.[3] For every country, we construct a region- and year-specific measure of weather shock. Given our focus on Colombia later in the paper, we present all the empirical patterns for west African countries and Colombia separately. After presenting such patterns, we discuss key characteristics of land markets in developing countries, which serve as the foundation of the model developed in Section 5.

2.1 Farm Size Dynamics and the Impact of Weather Shocks

Farm Size Dynamics We document two patterns related to farm size dynamics that parallel the empirical regularities observed in firm size dynamics—see for example Jaimovich et al. (2023) and Clementi and Palazzo (2016). Figure 1, Panel (a), shows that smaller farmers are more likely to exit farming.[4] The magnitudes are relevant: In Colombia, farmers operating below the average farm size are on average 75 percent more likely to exit (a share of 0.07 against a share of 0.04). In west African countries, farmers below the average farm size in their respective countries are 100 percent more likely to exit (a share of 0.16 against a share of 0.08). Panel (b) shows that smaller farms tend to grow, whereas larger farms tend to shrink.[5] In Colombia, farms below the average farm size grow, on average, 12 times, whereas

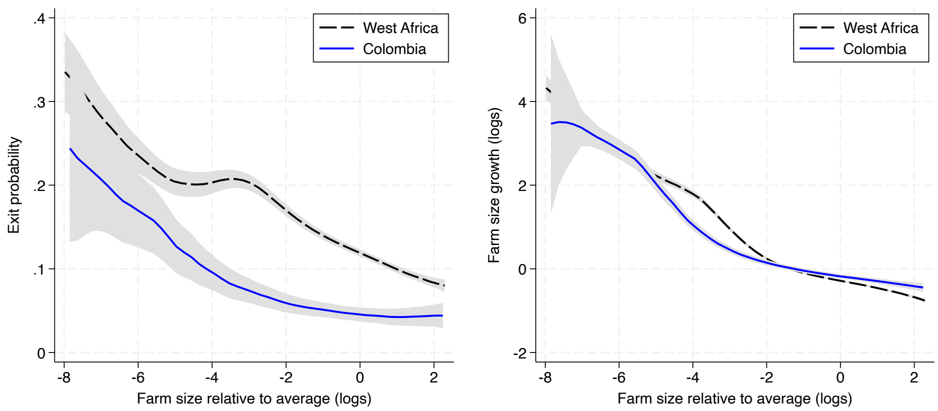
[3]ELCA stands for *Encuesta Longitudinal Colombiana de la Universidad de los Andes*, a longitudinal household survey with three rounds of data (2010, 2013, and 2016). EHCVM stands for *Enquête Harmonisée sur le Conditions de Vie des Ménages*, a nationally-representative household survey with two rounds of data (2017-2018 and 2021-2022) for Benin, Burkina Faso, Côte D'Ivoire, Guinea Bissau, Mali, Niger, Senegal, and Togo. We additionally use data for Nigeria from the LSMS General Household Survey-Panel.

[4]We define farm exit as a household owning no land. In the Colombian data, owning no land is strongly correlated with the probability of moving. Between 2010 and 2013, among farmers who sold all of their land by 2013, 18 percent moved to the urban area of the municipality compared to virtually 0 percent of farmers who did not sell their land. Similarly, among farmers who sold all of their land by 2013, 30 percent of them moved out of their region (to an urban area or another municipality), compared to 1 percent among farmers who did not sell their land.

[5]A small proportion of surviving farmers experience no change in their landholdings between surveys, given by a period of 3 years: about 6 percent of farmers show exactly no change in their landholdings, and about 11 percent show a change that is smaller than 10 percent.

Figure 1: Farm Size Dynamics across Countries

(a) Exit Probability (b) Farm Size Growth



Notes : This figure shows results from local polynomial regressions with bandwidth of 1. The data from west African countries are for 2018 and 2021 and the data for Colombia for 2010 and 2013. Exit probability equals 1 if the farmer holds no land in the second period. The farm size growth is the log of the ratio of farm size between the first and the second period. The x-axis is the log of the farm size divided by the average farm size in the respective country in the first period. Sample includes only farmers who own land in the first period.

farmers above the average grow by 5 percent. In west Africa, those below the average farm size grow around 13 times, whereas those above the average farm size shrink by 30 percent. We notice that the large growth among small farms comes, in part, from the extremely low basis: In west African countries, the average farm size is 3.45 hectares, whereas in Colombia the average farm size in our sample is 2.55.[6]

Weather Shocks. We estimate the impact of weather shocks on farm size dynamics using:

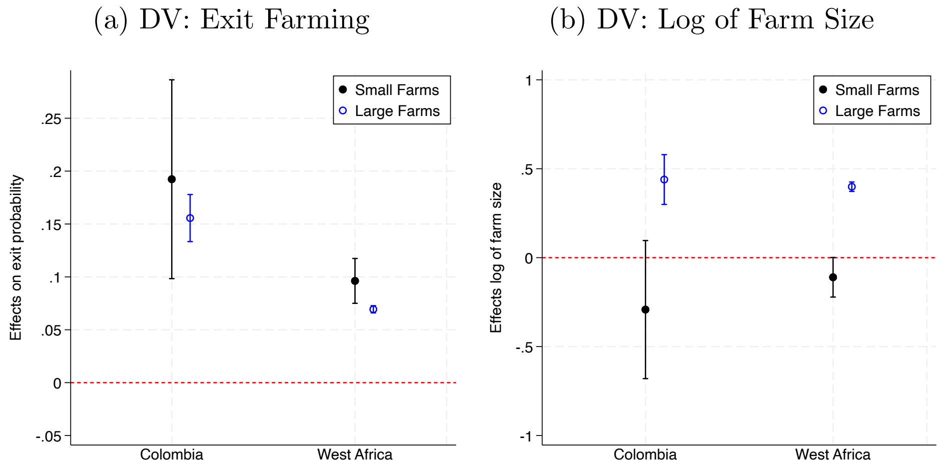
$$\mathrm{y}_{f,i,t} = \delta_t + \mu_f + \beta TempShocks_{i,t} + \gamma \left(TempShocks_{i,t} \times LF_f\right) + \epsilon_{f,i,t}, \tag{1}$$

where f indexes the farmer, i indexes the administrative region of farmer, and t is the year. y_{ft} is either (1) an indicator variable for whether a household exits farming and (2) the log of farm size. δ_t and μ_f are period and farmer fixed effects. $TempShocks_{i,t}$ is a measure of weather shock, which consists in the sum of atypical-temperature days (measured in hundreds) during the past two years—see details in Section 3—, LF_f is an indicator variable for whether the farm is above the average farm in its respective country, and ϵ_{ft} is the error term.

Figure 2 plots our estimates of equation (1). Panel (a) shows that weather shocks leads to a large increase the exit probability for all types of farmers, but with this impact being

[6]As discussed in Section 3, the farm-level data from Colombia, ELCA, focuses on smaller farms. In the agricultural census, the average farm size is 5.8 hectares.

Figure 2: Effects of Weather Shocks on Farmers' Exit and Farm Size Growth



Notes: This figure shows the impact of weather shocks on farmers' exit, defined as having no landholdings, and farmers' log of land area. We plot the point estimates of β and $\beta + \gamma$ from equation (1) and their respective 90% confidence interval. Standard errors clustered at the geographic administrative level in which the weather shock is measured. Sample size is 7,336 for Colombia and 50,742 for west Africa.

smaller for larger farms. Panel (b) shows the impact of weather shocks on the change in the log of farm size, which conditions on surviving farms. Small farms shrink after the weather shock, but large farms grow. The magnitudes of these effects are economically significant: an increase of one standard deviation in the number of days with atypical temperature over a two-year period increases the exit probability of smaller farmers in Colombia and west Africa by 3.9% and 3.8%, respectively.[7] Conditioning on surviving farms, an increase of one standard deviation in the weather shock measure reduces small-farm size in Colombia by 6%, but increases large-farm size by 9.2%. For the west Africa sample, an analogous increase in the severity of the weather shock reduces small-farm size 4.4% and increases large-farm size by 16.8%. The fact that large farms in average exit the market with the shock, whereas surviving large farms expand their operation, indicates the existence of general equilibrium effects where land relocates between households.

A final aspect we consider here is the role of new farmers. The EHCVM and LSMS household surveys provide insights into their significance: In 2021, 9 percent of all households with positive landholdings were landless in 2018. In Appendix Figure O.1, using data from west Africa, we find that weather shocks induce small farmers to exit, but they increase the probability that landless households become farmers. By design, ELCA focuses on households who owned land in the initial period, which makes this value, by construction,

[7] A one standard deviation increase corresponds to 20 additional days in Colombia, and 39 additional days in the west Africa sample. Estimation results for equation (1) in table form are shown in Table O.1 in the appendix.

close to zero.

2.2 Land Markets across Developing Countries

Three additional facts guide the interpretation of our reduced form findings and the design of our model. First, land rental markets in developing countries are in general thin, and farm operation usually coincides with land ownership. According to statistics from the National Agricultural Survey *Encuesta Nacional Agropecuaria* DANE (2019), about 85% of plots in Colombia were operated by owners in 2019, and 9% of farmers operated rented land. These figures are consistent with those reported by Acampora et al. (2022), who show that the proportion of agricultural households who rent out land across 6 African countries ranges between 0% and 5%, and stands in contrast to figures showing that the share of farmland operated under a rental agreement in the U.S. and the European union is about 40% and 46%, respectively.[8]

Second, the adoption of agricultural insurance is extremely low in developing countries. Unsubsidized agricultural insurance coverage rates in high-income countries average 41.7%, compared to 8% and 0.5% in lower-middle-income and low-income countries, respectively (Mahul and Stutley, 2008). In Colombia, according to (DANE, 2019), the share of farmers with any form of agricultural insurance is below 1 percent.

Third, Responses in the ELCA survey also suggest land distress-sales are not uncommon, and that farmers sell their land when experiencing financial distress to smooth consumption: 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. The use of land sales as a consumption smoothing device is also documented by Krishna (2010) in Africa.

These empirical regularities motivate the structure of our model. To capture the relationship between farm growth and farm size shown in Figure 1, we introduce idiosyncratic productivity shocks that induce farmers to expand or shrink their farm size. Additionally, farmers make occupational choices, in which they can sell their entire property to become workers. Crucially, these land and occupational choices are made in an economic environment with uninsured risk and no land rental markets. Building on the effects found in Figure 2, we simulate the effects of temporary aggregate productivity shocks, which affects all farmers in a region, to assess the relocations of land between agents and the aggregate implications to agricultural productivity.

Next, we turn our attention to Colombia, where we construct a more comprehensive

[8] U.S. Census of Agriculture (2022), and Eurostat (2020) figures.

longitudinal dataset on land transactions and farm size distribution. This dataset enables us to uncover a host of complementary facts that highlights the mechanisms driving the effects of weather shocks on the farm size distribution.

3 Data from Colombia

Our Colombian dataset is structured at two levels of disaggregation: municipality and farm. At the municipality level, it contains annual information on land transactions, farm size holdings, and weather shocks for 864 municipalities.[9] At the farm-level, it provides detailed information on farmers' consumption, occupational, and investment decisions. We provide a brief description of our datasets below, relegating details to Appendix OA. After presenting our data, we turn to the reduced-form effects of weather shocks in Section 4.

Land Transactions. We obtained administrative, property-level transaction information from the National Superintendence of Notaries (SNR)[10]—the government agency responsible for keeping records of all real estate market transactions. For each property-level transaction, we observe who is the buyer, who is the seller, and whether the transaction was a partial property sale, a full sale, or a mortgage.[11] These data cover all the properties that were once part of a *baldío*—a piece of land originally administered by the national government at some point allocated to a private individual. By 2014, these properties summed up to 23 million hectares—more than half of the currently privately-held land in the country (Sánchez and Villaveces, 2016; Arteaga et al., 2017). These data contain all transactions dating back to early 1900s, but we focus on the period between 2000 and 2011, which is when our other datasets are available. We construct a balanced yearly municipality-level data with the number of full sales, the number of partial sales, and the number of mortgage transactions. For complementary analyses, we split the number of sales into those involving buyers who previously owned land and those involving first-time buyers.[12]

Farm Size Distribution. Our farm size distribution data comes from the National Geographical Institute of Colombia (IGAC),[13] which has information on the universe of farm

[9] We exclude from our final sample of municipalities large metropolitan areas and municipalities with very few (i.e. below the 99^{th} percentile) properties registered in our farm size distribution data. Our final sample encompasses 85.3% of the rural population in the country.

[10] *Superintendencia de Notariado y Registro.*

[11] We do not observe the price nor the size of the properties that were sold.

[12] We proxy ownership by measuring whether a buyer ever appears in the SNR at the receiving end of a transaction, be it a sale, an inheritance or a government allocation.

[13] *Instituto Geográfico Agustin Codazzi.*

properties in Colombia.[14] We obtained access to municipality-level aggregate information from all properties 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, which is comparable to the area of Germany or California. The data consists in a yearly panel of municipalities with the number of farms, the number of owners, and average farm size for specific ranges of farm size. The data from the land registry is available for the period 2000-2011.[15]

Household-Level Data. We collect household panel data from the rural sample of ELCA, a survey conducted by the Universidad de Los Andes.[16] Here, we have 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 baseline sample includes 17 municipalities. We have detailed information on land ownership, consumption, and assets ownership. Moreover, we have information on migration of household members.

Weather Shocks. We construct our weather shock variables based on temperature information from ERA5 data, which is provided by the Copernicus Climate Change Service (C3S) of the European Centre for Medium-Range Weather Forecasts (ECMWF). This dataset contains global temperature information at the 0.25×0.25 degrees resolution ($\approx 28km^2$) on an hourly frequency since 1979. With this dataset, we first compute the historical quarterly distribution of daily average temperatures by considering all temperature measurements for municipality-day pairs within a calendar-quarter throughout the period 1979–2016. That gives us four temperature distributions per municipality, one per quarter. We then compute whether a municipality experienced an atypical temperature day if the temperature is below the 20$^{\text{th}}$, or above the 80$^{\text{th}}$ percentile of its corresponding quarterly temperature distribution. Indexing days by d, quarters by q, years by t, and municipalities by i, and denoting the 20$^{\text{th}}$ and 80$^{\text{th}}$ percentiles of quarterly temperature distributions as $\mu^{20}_{i,q(d)}$, and $\mu^{80}_{i,q(d)}$, we define an atypical temperature day in a municipality by

$$AtipicalDay_{i,t,d} = \mathbb{1}(\text{Temperature}_{i,t,d} \geq \mu^{80}_{i,q(d)}) + \mathbb{1}(\text{Temperature}_{i,t,d} \leq \mu^{20}_{i,q(d)}). \qquad (2)$$

[14]The data is designed to capture even informal properties that do not have all required formal documentation.

[15]Formally, the land registry records information on rural properties used for agriculture, which we consider analogous to farms. This is justified by the limited land rental markets, that create a strong link between property and farm boundaries. Supporting this notion, land registry data shows that, on average, each owner in a municipality reports only one property.

[16]Encuesta Longitudinal Colombiana.

To define a single shock variable for each municipality i in year t, we face two challenges. First, there is a potential lag between farmers' reaction to atypical days and the registration of land transactions or farm property in the data. In Colombia, it can take a few months—or even years—for land transactions to be recorded in the system. Second, the impact of the shock can influence farmers' decisions and outcomes with a lag—for example, atypical days experienced by a farmer in the end of a calendar year $t-1$ can have an effect on crop revenues only in t when the harvest season arrives. To address this potential mismatch between the impact of atypical days and economic outcomes, in our preferred specification we define the shock as the sum of atypical days lagged over the past two years:

$$TempShocks_{i,t} = \sum_{s=t-2}^{t-1} \sum_{d=1}^{365} AtypicalDay_{i,s,d}. \tag{3}$$

In addition to our preferred definition of a weather shock, because we compute for each municipality i the number of atypical temperature days accumulated within year t, we can experiment with alternative specifications in which we include the independent impact of atypical days within year t, together with its lags—see Appendix OB.1. Moreover, Appendix OB further reports results from a host of alternative definitions of our temperature shock, including three alternative definitions of what constitutes an atypical day, different formulations of a single shock variable with accumulations over different lags of years, and an entirely different source of data on weather shocks based on SPEI —which is preferred by some papers in the literature. Reassuringly, our qualitative results remain largely unaffected by these different approaches.

4 Reduced-Form Impact of Weather Shocks

We organize this section as follows. We first present our municipality-level regressions in Section 4.1, which shows how land transactions and the farm size distribution react to weather shocks. We then present our household-level regressions in Section 4.2. Section 4.3 discusses potential mechanisms. Throughout our exposition, for brevity, we focus on our preferred specification. Section 4.4 discusses in detail all the alternative specifications that we employ.

4.1 Municipality-Level Regressions

We estimate the impact of temperature shocks on land transactions and farm size using:

$$s_{i,t} = \beta TempShocks_{i,t} + X'_{i,t}\delta + \eta_i + \kappa_t + \varepsilon_{i,t}. \tag{4}$$

Here, $s_{i,t}$ represents the number of land transactions (sales or mortgages), the log number of farm owners, or the log average and median farm size in municipality-year pair i, t. $X_{i,t}$ is a vector that includes time-varying municipality characteristics. Specifically, rainfall at t and previous lags, the cumulative number of land allocations in the municipality up to t, which controls for the land on which we observe transactions, a dummy variable for cadastral updates, and the total municipal land area recorded in the registry, which accounts for changes in registry coverage. The model also includes municipality fixed effects η_i to control for time-invariant heterogeneity across municipalities, and year fixed effects κ_t to control for time-specific shocks to land markets and farm size common to all municipalities. Robust standard errors in all regressions are clustered at the municipality level.

Equation (4), as well as all subsequent specifications, rely on the identifying assumption that, conditional on the fixed effects and the set of control variables, there are no municipality-specific, time-varying unobservables correlated with the occurrence of atypical temperature shocks. This is a standard assumption in the literature (see, e.g., Dell et al. (2014)). Section OB.2 shows that our results are robust to specifications that include state-specific time trends, additional control variables, and alternative measures of atypical temperature.

Table 1: Temperature Shocks and Land Transactions

	Total	Full	Partial	Mortages
	(1)	(2)	(3)	(4)
TempShocks	2.537***	2.013***	0.523**	1.046***
	(0.536)	(0.504)	(0.229)	(0.237)
Observations	10,021	10,021	10,021	10,021
R^2	0.905	0.902	0.636	0.757
Mean Dep. Var	12.62	10.83	1.79	2.62

Notes: Data from the National Superintendency of Notaries (SNR) records. Coefficient estimates from equation (4). The dependent variable in column 1 is total number of land sales (full + partial), in column 2 the total number of full sales, in column 3 the number of partial sales, and in column 4 the number of land mortgages. The main independent variable is the total number of atypical temperature days in the past two years ($t-1$, $t-2$) divided by 100. All regressions include municipality and year fixed effects as well as cumulative rainfall and its 5 lags, number of land allocations, area covered by the land registry, and an indicator of registry updates. See texts for more details. See texts for more details. Robust standard errors clustered at the municipality level reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Effects on Land Transactions. Table 1 presents the OLS estimates from equation (4) on four measures of land transactions. Column 1 shows results for the total number of land sales. Columns 2 and 3 separate this into 'full' sales (entire property transfers) and 'partial' sales (fractional property transfers). Column 4 presents the number of land mortgages. Consistent with the results from Figure O.9, columns 1 and 4 show that increases in the days with atypical temperature induce land sales and mortgages. In particular, an increase of 100 (roughly 2 standard deviations) in the number of days with atypical temperature over the previous two years increases the number of land sales by 20%. Columns 2 and 3 show that this increase is driven by an increase of 19% in full sales and of 29% in partial sales. Column 4 shows that additional 100 days of atypical temperature in the municipality increase by 38% the number of mortgages taken out by farmers against their properties.

Effects on Farm Size and Number of Owners. Table 2 presents the results of estimating equation (4) on the number of owners and our measures of farm size.[17] More days of atypical temperature in a municipality during the previous two years lead to an increase in the number of landowners (column 1), and to lower average and median farm size (columns 2 and 3). The coefficient estimates in this table suggest that an additional 100 days of atypical temperature over a two-year period increases the number of landowners by 1.5% and reduces median farm size by 2.1%.

Table 2: Temperature Shocks and Average Farm Size

	Number of Owners (1)	Mean Farm Size (2)	Median Farm Size (3)
$TempShocks$	0.015***	-0.015***	-0.021*
	(0.005)	(0.005)	(0.012)
Observations	10,021	10,021	10,021
R^2	0.992	0.993	0.974
Mean Dep. Var	2,501.13	32.06	16.15

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 ($t-1$, $t-2$) divided by 100. All regressions include municipality and year fixed effects as well as log cumulative rainfall and its 5 lags, log number of land allocations, log area covered by the land registry, and an indicator of registry updates. See texts for more details. Dependent variables are the log number of land owners in the municipality (column 1), log average farm size (column 2), log median farm size (column 3). *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.

Effects on Landless Buyers. The increase in the number of landowners per municipality shown in Table 2 could come from landless households or from farmers in neighboring municipalities purchasing land. To better understand which of these two sources drives the effect, we exploit the fact that our transaction data has information on both the buyer and the seller participating in each observed land sale. We use this to build the list of individuals who own any land within each *departamento* (the administrative unit above the municipality, for which we have 28 units) at the start of every year, and match it to the list of land buyers in each municipality throughout the following two years. This allows us to determine if buyers already owned land in any of the municipalities within the departamento or not.

[17] Farm size is defined based on the registered size of the plots. Since some farmers might own more than one plot, the average farm size tends to be different from the ratio of land to the number of owners.

Details on the construction of these lists the matching process between them, and alternative approaches for data construction is discussed in Appendix OB.3.

We decompose the estimated impact of weather shocks on land sales shown in Table 1 into sales made to landless buyers and sales made to farmers who were already landowners by the time of the sale. Table 3 shows that extreme temperature increases land sales to both types of buyers, but that effects are larger among landless owners, who on average make up 74% of the weather-driven land sales.

Table 3: Effect of Temperature Shock on Land sales by Type of Buyer

	(1) Total Sales	(2) Sales to Landless Buyers	(3) Sales to Already Owners
$TempShock_{i,t}$	2.537***	1.866***	0.671***
	(0.536)	(0.408)	(0.192)
Observations	10,021	10,021	10,021
R^2	0.905	0.889	0.838
Mean Dep. Var	12.62	9.37	3.25

Notes: Data from the National Superintendency of Notaries (SNR) records. Coefficient estimates from equation (4). The main independent variable is the total number of atypical temperature days in the past two years ($t-1$, $t-2$) divided by 100. Controls are accumulated allocations, area covered by the land registry, an indicator of registry updates, accumulated precipitation and five lags. Robust standard errors clustered at the municipality level reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Effects across Initial Farm Size Groups. Table 2 does not provide information on the initial size of the farms that are being fragmented. A reduction in the average farm size in a municipality could arise from the division of large estates into medium-sized properties without any change in the number of small farms, or it could result from the fragmentation of small farms into even smaller parcels, without affecting the number of larger properties. To investigate this issue, we estimate the changes in the number of owners within fixed farm-size bins over time. First, we split the distribution of farm size in each municipality into quantiles, ensuring that each quantile has the same number of owners in the initial year of our sample.[18] We then compute the number of owners within each of these fixed size bins for each subsequent year.

Denote as $\{q_i^1, ..., q_i^J\}$ the areas defining each of the j quantiles of the farm size distribution in municipality i in the initial year of our sample, and denote as $AreaOwned_{o,i,t}$ the total landholdings of owner o in municipality i and year t. We compute for each year the number

[18]We take the initial distribution as the year 2000, for which 97% of municipalities have registry information. For the remaining municipalities, we use the first year in which they appear in the land registry dataset.

of owners with total landholdings within each of these fixed size bins as:

$$NumOwners^j_{i,t} \equiv \sum_{o \in i} \mathbb{1} \cdot [AreaOwned_{o,i,t} \in (q_i^{j-1}, q_i^j)], \tag{5}$$

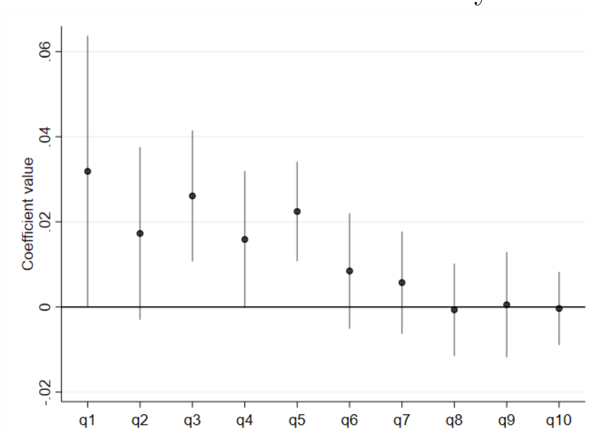
where $j = 1, ..., J$, and $q_i^0 = 0$ for all i. We use this variable to estimate independent regressions (one per quantile j) of the form:

$$NumOwners^j_{i,t} = \gamma^j TempShocks_{i,t} + X'_{i,y}\xi^j + \mu_i^j + \kappa_i^j + \omega^j_{i,t}, \tag{6}$$

where all the right-hand-side variables are the same as in (4)

Figure 3 reports our estimates. Extreme temperature shocks cause a sizable increase in the number of owners with farms on the lower 5 deciles of the initial distribution, but close to zero and not statistically significant effects on the number of owners in the 5 top deciles.[19] The number of large landowners remains stable while the number of small owners, suggesting that the decline in average farm size is driven by new owners acquiring parcels of land subdivided from small estates, instead of small farmers being solely bought by larger farms.

Figure 3: Temperature Shocks and Number of Owners by Initial Distribution Quantiles



Notes: OLS estimates of the γ coefficients according to equation (6), 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 ($t-1$, $t-2$) divided by 100. The dependent variable is the log number of owners per quantile. Controls are log accumulated allocations, log area covered by the land registry, an indicator of registry updates, log accumulated precipitation and five lags. Regressions also include year and municipality fixed effects. Error bars display 95% confidence intervals for standard errors clustered at the municipality level.

[19] Regression results in table form are in O.11 in the appendix. Appendix Figure O.11 shows analogous estimations for alternative partitions ($j = 5$, and $j = 20$) of the initial farm size distribution.

4.2 Household-Level Regressions

We now examine whether the effects of temperature shocks on household outcomes align with the aggregate patterns in land sales and farm size, and support our interpretation of land sales as a consumption-smoothing device. Similarly to Section 2, we estimate:

$$y_{f,i,t} = \alpha TempShocks_{i,t} + X'_{i,t}\tau + \iota_f + \kappa_t + \psi_{f,i,t}, \quad (7)$$

where $y_{f,i,t}$ denotes an outcome for household f in municipality i, and survey round $t = \{2010, 2013, 2016\}$. Vector $X_{i,t}$ denotes controls for log aggregate rainfall at t and five lags, ι_f represents household-level fixed effects, and κ_t year fixed effects.

Table 4 shows that temperature shocks decrease the likelihood that households own land and increase the probability of landowners having less than three hectares (columns 1 and 2). Additionally, there is an imprecisely estimated negative coefficient for log farm size—but as shown earlier, effects are heterogeneous across small and large farms. Temperature shocks also increase that households sell other types of assets: as shown in columns 4 and 5, farm animals and the value of a principal component index of asset ownership, which includes appliances, vehicles and other household assets, fall with the temperature shock. Appendix Table O.8 shows that this is driven by a decrease in the likelihood that households own a wide variety of assets. Column 6 shows that households are also more likely to migrate when experiencing the shock, a result that is consistent with findings from Ibáñez et al. (2022) for El Salvador. Lastly, temperature shocks have a sizable effect on the monetary value of per-capita consumption –a 12% drop per 100 additional days according to column 7–, indicating that households are unable to fully smooth consumption despite the various types of asset liquidation that we observe.

Table 4: Temperature Shocks and Household Decisions

	Household Has Land (1)	Farm Size $\leq$ 3 ha (2)	Farm Size (3)	HH sold Animals (4)	Asset Index (5)	Household Migrated (6)	Consumption per capita (7)
TempShocks	-0.065***	0.055***	-0.074	0.043***	-0.118***	0.062***	-0.118***
	(0.016)	(0.019)	(0.068)	(0.008)	(0.014)	(0.017)	(0.023)
Observations	11,905	9,918	9,918	11,422	11,422	11,905	11,418
R^2	0.639	0.718	0.741	0.931	0.611	0.533	0.731
Mean Dep. Var	0.88	0.74	2.96	0.69	-0.00	0.12	2.66

Notes: Data from ELCA. Dependent variables from left to right: a dummy indicating if household owns any land; a dummy indicating if household's landholdings are below 3 hectares; log farm size; a dummy indicating if household sold animals; a principal components asset index that include household and farm assets, a dummy indicating if household migrated between survey waves; and the log value of per capita consumption in 2016 Colombian pesos (in millions). All regressions include a control of log aggregate rainfall at t and five lags, and household and time 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.

4.3 Mechanisms

Section 5 develops a theoretical model that rationalizes our collection of results through the lack of access to insurance mechanisms. Farmers sell land and exit agriculture, in part, to smooth consumption when experiencing negative income shocks. In particular, small farmers are more likely to use these coping strategies because of their lower consumption levels, which implies higher marginal utility of consumption and greater gains from increasing present versus future consumption. Because the shock is temporary, the future value of becoming a farmer for landless households remains unaffected, but the price of land that landless households face still falls, because current farmers who are experiencing the negative shock are trying to sell their land to smooth consumption. Quantitatively, that can lead to an increase in the number of farmers and a decrease in average farm size after the shock.

We engage with three alternative mechanisms that could rationalize some of our results. First, the lack of contiguity between large and small farms might limit the expansion of larger farms after the shocks. To assess this hypothesis, we construct a measure of contiguity between large and small farms in each municipality using land registry maps from 2017. Within each municipality, we calculate the share of farms below the 10th percentile of the size distribution that are contiguous to at least one farm above the 90th percentile. We then classify municipalities with high-contiguity as those with a share above the national median. Second, small farms might be more easily converted to residential or recreational purposes, which is more likely in the outskirts of cities due to urban expansion. We investigate this hypothesis by focusing on municipalities located farther than the median distance from main cities, where these alternative uses are more frequent.[20]

Table 5 shows that the evidence on the importance of these two alternative mechanisms is weak. Panel (a) includes an interaction term with an indicator for high contiguity. Column 3 provides some evidence that the lack of contiguity may interact with the shock, but coefficients are small and not statistically significant for the number of owners and average farm size. Moreover, Appendix Table O.10 shows that the result on median farm size is not robust to an alternative measure of contiguity based on the GPS coordinates from the 2014 National Agricultural Census. Panel (b) adds an interaction term with a dummy variable for municipalities located farther than the median distance from the main cities. Results are statistically insignificant for columns 2 and 3 and, contrary to what this mechanism would imply, we observe a larger increase in the number of landowners in municipalities farther from large cities.

[20] Appendix OB.4 also discuss the land ceiling regulation that applies in Colombia for government allocated land and present some results showing that the restrictions to land consolidation imposed by these size ceilings are not likely to explain land fragmentation after extreme temperature.

For our third alternative mechanism, we investigate if our results are due to institutional factors stemming from Colombia's land regulation policies. Law 160 of 1994 imposed municipality-specific land ceilings that placed a cap on the amount of government-allocated land that any private individual can buy. Appendix OB.4 re-estimates equation 4 and include an additional interaction term for municipalities with an above-the-median share of government-allocated farmland—i.e., where this restriction would be most binding—and show that the land ceiling policy is not driving our results.

Table 5: Temperature Shocks and Farm Size - Heterogeneous Effects

	Number of Owners	Mean Farm Size	Median Farm Size
	(1)	(2)	(3)
	Panel A: Land Registry Map - Contiguous Plots		
$TempShocks$	0.012**	-0.014**	-0.032**
	(0.005)	(0.006)	(0.014)
$TempShocks \times High$	-0.004	0.004	0.024*
	(0.005)	(0.006)	(0.013)
Observations	9,499	9,499	9,499
R^2	0.991	0.993	0.973
Mean Dep. Var	2,573.13	32.05	16.23
	Panel B: Distance to markets		
$TempShocks$	0.010*	-0.013**	-0.027*
	(0.006)	(0.006)	(0.015)
$TempShocks \times High$	0.011*	-0.006	0.007
	(0.006)	(0.006)	(0.012)
Observations	9,683	9,683	9,683
R^2	0.991	0.993	0.974
Mean Dep. Var	2,566.22	31.98	16.10

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 ($t-1$, $t-2$) divided by 100. Controls are , log accumulated precipitation during and five lags, log accumulated allocations, log area covered by the land registry and an indicator of registry updates. Regressions also include year and geographic fixed effects. *High* in panels A indicates a dummy variable equal to one for municipalities with a high (above the median) share of farms below the 10th percentile of the size distribution that are contiguous to at least one farm above the 90th in the 2017 land registry map. In panel B it indicates a dummy variable equal to one for municipalities located at a diving distance to the nearest city above the national media. See text for more details. Standard errors clustered at the municipality level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

4.4 Alternative Specifications and Measures of Weather Shocks

In Appendix OB we show that our reduced-form results are robust to a host of alternative specifications and specific variable definitions. We first show that results for the impact of weather shocks on land sales, mortgages, number of owners, and mean and median farm size are robust to a flexible-lag specification that allows for variation in the timing of the effects with respect to the onset of the shock. This specification—described in equation (O.1)—estimates independent coefficients for the impact of weather shocks occurring each year between t and $t-7$. Results for this exercise are presented in Figures O.9 andO.10, and show that extreme temperature events can have lasting effects several years after the shock takes place and that our results do not depend on the specific choice of lags used for the definition of the shock.

Second, Tables O.4 and O.5 report estimates from equation (4) using alternative definitions of temperature shocks. Panels A and B define atypical temperature using as thresholds the 5th and 95th percentiles and the mean $\pm$ 1.5 standard deviations of the temperature distribution, respectively. Panel C uses the same threshold as the main specification, but adjusts the time reference for computing the temperature distributions to 1990-2011. Panels D and E define shocks using the SPEI index.

Finally, Tables O.6 and O.7 provide results from alternative specifications. Panel A incorporates *departamento*-specific time trends, which allows us to account for potential spurious correlations between regional temperature shock trends and our variables of interest. Panels B and C control for forced displacement and homicide rates to capture the impact of violent conflict within municipalities. Panel D presents results from the main specification, clustering standard errors by municipality and departamento $\times$ year.

5 A Quantitative Model of Farm Size Dynamics

Motivated by the previous empirical facts about the dynamics of farm size distribution and the influence of weather shocks, in this section we develop a dynamic, heterogeneous household model in which uninsured households choose whether to be a farmer or a worker, and, if they choose to be a farmer, how much land to buy for production. Based on our institutional setting, our model features no rental markets for land. Instead, households have to buy their land in every season, prior to the harvest season, under the uncertainty about their future farm productivity. As such, households purchase land to build wealth, but land also acts as a smooth consumption device. The joint occupational and landholding decisions of heterogeneous households determine the farm size distribution of the economy.

After laying out the model, we describe our calibration procedure. In the next section, we use our calibrated model to quantify the aggregate implications of farm size dynamics and weather shocks to aggregate agricultural productivity.

5.1 Environment

The economy operates over time t, which is discrete. Each period represents a year, during which one harvest season occurs. The economy is endowed with an exogenous supply of land

L and of households N.

There are two occupations in this economy: farmers and workers. Farmers own land $\ell_t > 0$, have farming productivity s_t, and earn agricultural profits π_t. Their production technology is given by

$$y_t = (Z_t s_t)^{1-\gamma} \left(\ell_t^{\alpha} k_t^{1-\alpha}\right)^{\gamma}, \tag{8}$$

where Z_t is the aggregate climatic condition of the economy and k_t is the farm capital. $\gamma \in (0,1)$ is a span-of-control term and $\alpha \in (0,1)$ is the share of land. The rental rate of farm capital, r, is exogenous and independent of t. Optimal choice of farm capital gives the following equation for agricultural profits

$$\pi_t \equiv \pi(Z_t, s_t, \ell_t) = \xi \left[(Z_t s_t)^{1-\gamma} (\ell_t^{\alpha})^{\gamma}\right]^{\frac{1}{1-\gamma(1-\alpha)}}, \tag{9}$$

where ξ is a constant that incorporates the production function parameters and the rental price of capital.[21]

Workers are landless $\ell_t = 0$, receive signals of their farming productivity s_t, and earn wages w_t, which are not affected by the climate condition Z_t.

At each period t, households receive the option of becoming a farmer in $t+1$ with probability $1-\delta$, otherwise they become workers—δ is therefore an exogenous exit probability for farmers. If given the option of becoming a farmer, households draw a taste shock for each occupation, denoted by ε_F for farming and ϵ_W for working, and decide whether to exercise that option. Should they choose to become a farmer, they must purchase land ℓ_{t+1} for the next harvest season. In any other case, they become workers in $t+1$, sell any land they might own, and consume all of their earnings.

The timing of the model within each period t is as follows:

1. The aggregate climatic condition Z_t is observed.

2. Workers observe their farm productivity s_t and their wage w_t.

3. Households receive the option of becoming a farmer and choose their occupation for $t+1$.

4. Farmers receive π_t, workers earn w_t, land market opens, and households consume.

5. Farmers transition to their next farming productivity s_{t+1}.

[21] Here, $\xi \equiv [r]^{\frac{\gamma(1-\alpha)}{1-\gamma(1-\alpha)}} \left[\gamma(1-\alpha)^{\frac{\gamma(1-\alpha)}{1-\gamma(1-\alpha)}} - \gamma(1-\alpha)^{\frac{1}{1-\gamma(1-\alpha)}}\right]$.

5.2 Distributional Assumptions

Farming productivity s_t follows a stochastic process $\log s_{t+1} = \rho \log s_t + \sigma \epsilon_{t+1}$, where $\epsilon_t \sim \mathcal{N}(0,1)$ is a productivity shock, $\rho \in (0,1)$ is the persistence, and $\sigma > 0$ is the volatility of ϵ_t. We denote the conditional distribution of farming productivity is $G_s(s_{t+1}|s_t)$, with density $g(s_{t+1}|s_t)$. Workers draw their signal of farming productivity from the stationary distribution of s_t from $\tilde{G}_s(s_t)$, with density $\tilde{g}_s(s_t)$. The wage distribution of workers comes, given by $G_w \sim \log \mathcal{N}(\mu_w, \sigma_w)$ with density g_w. Taste shocks ε_F and ϵ_W are drawn independently from an extreme value type one distribution (EVT1) with dispersion parameter κ.

5.3 Optimization problem

Households live forever and are indexed by ω_t. Their preference is described by their long-life utility $U(c) = \mathbb{E}\left[\sum_0^\infty \beta^t u(c_t)\right]$, where c_t is consumption and $u(c_t)$ is the utility function—which is assumed to be log utility. Households are characterized by a triple (w_t, s_t, ℓ_t). Notice that farmers earn no wage, so that $w_t = 0$, and workers own no land, so that $\ell_t = 0$ and $\pi_t = 0$.

The households' optimization problem can be written in a recursive form. They choose their occupation and land purchases based on the start-of-the-period value function

$$V(\omega_t) = (1-\delta) \max\left\{V_F(\omega_t) + \varepsilon_F, V_W(\omega_t) + \epsilon_W\right\} + \delta V_W(\omega_t) \tag{10}$$

where the value of becoming a farmer in $t+1$ in the current period t, $V_F(\omega_t)$, is

$$\begin{aligned} V_F(\omega_t) &= \max_{c_t, \ell_{t+1}} \left\{u(c_t) + \beta \mathbb{E}_s\left(\mathbb{E}_\varepsilon\left(V(\omega_{t+1})\right)|s_t\right)\right\} \\ \text{s.t. } c_t &= w_t + \pi(Z_t, \omega_t) + p_t(\ell_t - \ell_{t+1}), \end{aligned} \tag{11}$$

and the value of becoming a worker in $t+1$ in the current period t, $V_W(\omega_t)$, is

$$\begin{aligned} V_W(\omega_t) &= \max_{c_t} \left\{u(c_t) + \beta V_0(\omega_t)\right\} \\ \text{s.t. } c_t &= w_t + \pi(Z_t, \omega_t) + p_t \ell_t \end{aligned} \tag{12}$$

Here, $V_0(w_t)$ is an exogenous value of becoming a worker. For simplicity, we assume that $V_0 = V_0^W$ if $w_t > 0$ and $V_0 = V_0^F$ if $w_t = 0$. This formulation captures, in a parsimonious way, the fact that farmers and workers might perceive different values of becoming a worker, for example due to sectoral relocation costs.

The optimal choice of land in equation (11) is $\ell^*(\omega_t)$. When farmers do not buy any additional land, so that $\ell_t = \ell^*(\omega_t)$, consumption equals agricultural profits $c_t = \pi(Z_t, \omega_t)$. In contrast, if $\ell^*_{t+1}(\omega_t) > \ell_t$, then farmers sacrifice consumption in t to increase potential profits in the future. Lastly, if $\ell^*_{t+1}(\omega_t) < \ell_t$ then farmers increase their consumption today at the expense of expected future profits.

Because ε_F and ϵ_W are drawn from a EVT1 distribution, as is common in the literature, the expected value is given by the following equation

$$\mathbb{E}_\varepsilon (V(\omega_t)) = \kappa \log\left(\exp\left(\frac{1}{\kappa} V_F(\omega_t)\right) + \exp\left(\frac{1}{\kappa} V_W(\omega_t)\right)\right), \tag{13}$$

and the probability that a household ω_t will choose to farm in $t+1$ if the option arises is

$$\mu(\omega_t) = \frac{\exp\left(\frac{1}{\kappa} V_F(\omega_t)\right)}{\exp\left(\frac{1}{\kappa} \mathbb{E}_\varepsilon (V(\omega_t))\right)}. \tag{14}$$

5.4 Evolution of distributions

At the beginning of period t, the density of households is $g(\omega_t)$ and the cumulative distribution is $G(\omega_t)$. The end-of-the-period density of households who become farmers in $t+1$ is

$$h_F(\omega_t) = (1-\delta)\, \mu(\omega_t)\, g(\omega_t),$$

and the distribution of households who become farmers in $t+1$ is

$$g(\omega_{t+1} | w_{t+1} = 0, \ell_{t+1} > 0) = \int g(s_{t+1}|\omega_t)\, a(\ell_{t+1}|\omega_t)\, h_F(\omega_t)\, d\omega_t, \tag{15}$$

where $a(\omega_t)$ is a policy function that equals one if household ω_t chooses $\ell_{t+1} = \ell^*(\omega_t)$ and zero otherwise. The end-of-the-period density of households who become workers in $t+1$ is

$$h_W(\omega_t) = \delta \mu(\omega_t)\, g(\omega_t),$$

and the distribution of households who become workers in $t+1$ is

$$g(\omega_{t+1} | w_{t+1} > 0, \ell_{t+1} = 0) = \int \tilde{g}_s(s_{t+1})\, g_w(w_{t+1})\, h_W(\omega_t)\, d\omega_t. \tag{16}$$

The density of households in $t+1$ is therefore the sum of conditional densities in equations (15) and (16).

5.5 Land market clearing

The end-of-the-period demand for land in period t is

$$L_t^D = \int \ell^* (\omega_t) \, h_F (\omega_t) \, N d\omega_t. \tag{17}$$

In each period t, the price of land p_t ensures that land market clears

$$L = L_t^D. \tag{18}$$

5.6 Stationary Equilibrium

Given distributions $\{G_s, \tilde{G}_s, G_w\}$, a mass of households $\{N\}$, a land endowment $\{L\}$, production function parameters $\{\alpha, \gamma\}$, an exogenous exit probability $\{\delta\}$, an exogenous farm capital rent $\{r\}$, and an aggregate climatic condition $\{Z^*\}$, a stationary competitive equilibrium are optimal consumption and land choices $\{\ell^*,c^*\}$, a land price $\{p^*\}$, and an invariant distribution $\{G^*\}$ such that: land and consumption choices are optimal and satisfy (10) to (12), occupational choices given in equation (14) are optimal (14), land markets clear so that equation (17) is satisfied, and the stationary distribution is consistent with equations (15) and (16).

5.7 Calibration

To simulate the model, we need parameter values for the agricultural production function (α, γ, ρ, σ, and Z_t), the wage of workers (μ_w and σ_w), the occupational choices (κ, δ, V_0^W and V_0^F), the discount rate (β), and the labor and land endowments (N and L^S). Our procedure combines estimation with a calibration algorithm. Several parameters we pick from the literature, the remaining ones are calibrated so that the stationary equilibrium of the model matches a series of statistics in the data related to farm size distribution and households occupational choices. Table 6 summarizes the parameters and the statistics that we target in our calibration.

Before applying our calibration algorithm, we pick from the literature or estimate the following parameters. First, for the agricultural production function, we set $\gamma = 0.47$ and $\alpha = 0.46$ based on estimates from Avila and Evenson (2010) for Colombia. These values imply a share of land of 0.22, which is largely within the values used in the literature. The discount rate comes from Greenwood et al. (2019), who estimate this parameter in the context of a developing country. We borrow κ from the literature on migration and sectoral relocation. Here, the literature finds values between 0.5 and 4, by using regional and sectoral

Table 6: Summary of Calibration of Parameters

Symbol	Value	Description	Target/Source	Target Value
γ	0.47	Share of inputs	Avila and Evenson (2010)	-
α	0.22	Share of land	Avila and Evenson (2010)	-
β	0.96	Discount rate	Greenwood et al. (2019)	-
κ	4	Occupational choice elasticity	Migration Literature	-
ρ	0.75	Serial correlation in yields	Regression of yields	-
N	0.17	Total mass of agents	Average Farm Size	5.98
σ	2.2	Volatility of shocks	S.d. of farm size	15.50
δ	0.05	Exogenous exit probability	Share of large farms who exit	0.05
V_0^F	-51.1	Value of becoming a worker	Share of small farms who exit	0.08
V_0^W	20.3	Value of staying as a worker	Share of workers in agriculture	0.15
μ_w	-2.37	Average of log wage	Mean farm size of new entrants	1.32
σ_w	1.12	S.d. of log wage	S.d. of farm size of new entrants	2.39
Z_{ss}	1	SS aggregate productivity	Normalization	-
L^S	1	Total supply of land	Normalization	-

flows of workers. We set $\kappa = 4$, so that we capture a lower sectoral elasticity, which is consistent with higher frequency of our data (annual). Lastly, we estimate ρ using farm-level longitudinal data from ELCA, by estimating the autocorrelation over the years between model-implied measures of farm productivity—see details in Appendix OC. Without loss of generality, we normalize $Z_{ss} = 1$, the steady-state aggregate productivity, and $L^S = 1$.

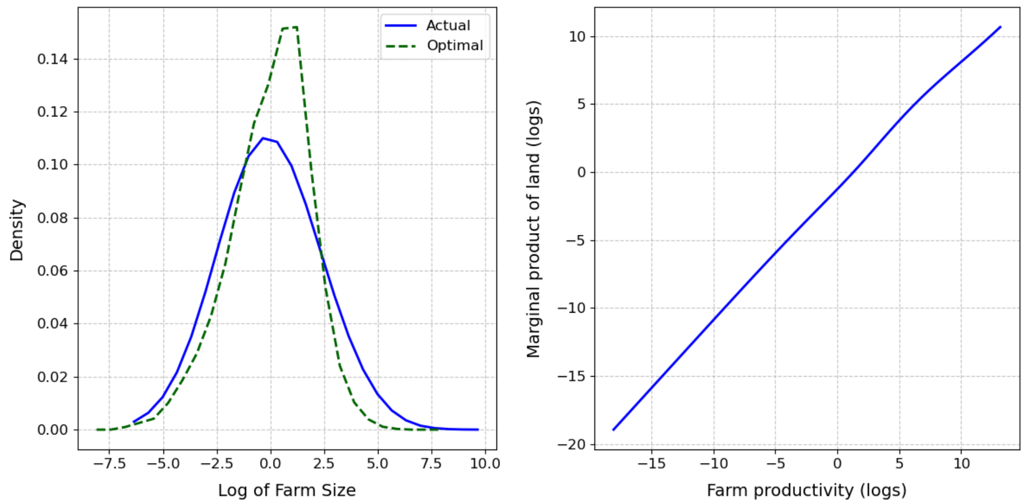
We calibrate the remaining parameters $\{N, \sigma, V_0^F, \delta, V_0^W, \mu_w, \sigma_w\}$ so that the stationary equilibrium matches six statistics in the data exactly. Specifically, N matches the average farm size, σ the standard deviation of farm size distribution, V_0^F the exit rate of small farms (below 2.7 hectares), and δ the exit rate of larger farms (above 2.7 hectares), V_0^W, the share of farmers in the economy, μ_w the average farm size of new entrants, and σ_w the standard deviation of the farm size distribution of new entrants. The farm size distribution statistics are constructed using data from the agricultural census of 2014. The farm size distribution for new farmers and the exit rates are computed using our longitudinal farm level data (ELCA). The share of workers employed in agriculture comes from ILO.

6 Farm Size Dynamics and Agricultural Productivity

This section uses our calibrated model to assess: (1) how risk shapes agricultural productivity and misallocation in the stationary equilibrium; and (2) how weather shocks affect transition dynamics in terms of the farm size distribution and agricultural productivity.

Figure 4: Farm Size Distribution and Misallocation in the Stationary Equilibrium

(a) Farm Size Distribution (b) Misallocation



Notes: This figure shows the farm size distribution generated by the model and the implied farm size distribution if the marginal product of land equalized across farmers, for example, due to the existence of rental markets. The optimal farm size is computed based on equation (19), assuming the stationary distribution of farm productivity.

6.1 Risk and Agricultural Productivity in the Stationary Equilibrium

We start by examining the implications of uninsured risk to misallocation, as measured by the dispersion in marginal product of land based on the stationary equilibrium. If markets were operating perfectly, land would be allocated to farmers with the highest marginal product of land, until the marginal product of land were equalized across farmers. In our model, however, two mechanisms prevent the marginal product of land from being equalized across farmers: (1) farmers have to buy their land *before* observing their effective farm productivity, so that we have a mismatch between the period in which the factor is chosen and the actual realization of the productivity; and (2) the lack of credit markets will induce farmers to treat land transactions as a consumption smoothing device.

Let us start by defining the marginal product of land (MPL) in the context of our model, given optimal choices of farm capital. Using equation (9), the marginal product of land of a farmer is

$$\mathrm{MPL}_t \propto \frac{\pi_t}{\ell_t}. \tag{19}$$

If there were a rental market for land, then the marginal product of land would equalize across farmers, via land rents. In that case, the land allocation across farmers would be proportional to farm productivity s_t.[22]

Figure 4(a) shows the optimal and the actual farm size distribution. The optimal dispersion in farm size is substantially smaller, with fewer small and large farms. Specifically, the standard deviation of the actual farm size distribution is 29 percent higher than the standard deviation of the optimal farm size distribution. Panel (b) shows that more productive farms have a larger marginal product of land—similarly to findings in Adamopoulos et al. (2022). Therefore, the optimal farm size distribution would require a relocation of land from less to more productive farms. To understand the implications of this misallocation to output, the total output of the economy in the baseline calibration if land were allocated optimally is 32 percent higher than the actual output.

6.2 Weather Shocks and Transition Dynamics

We now turn to the transition dynamics. We simulate a temporary, unexpected shock of 25 percent reduction in the aggregate agricultural productivity Z_t, with a low correlation over time, such that Z_t is virtually back to its level pre-shock level by the fifth period.[23] Figure 6 (a) shows the impact on land prices. They fall by 20 percent with the climate shock, and remain. As a result, agricultural output falls by approximately 25 percent and farmers' consumption falls by 15 percent for those who exit, and 26 percent for those who stay. Here, consumption smoothing occurs mostly via changes in occupational choice—farmers who do not exit sacrifice consumption to keep their landholdings. Figure 5(a) shows that smaller and less productive farmers are more likely to leave the agricultural sector, since they have a higher utility from consumption — Appendix OD derives analytical results that show how productivity and price shocks affect occupational choice of farmers with different landholdings.

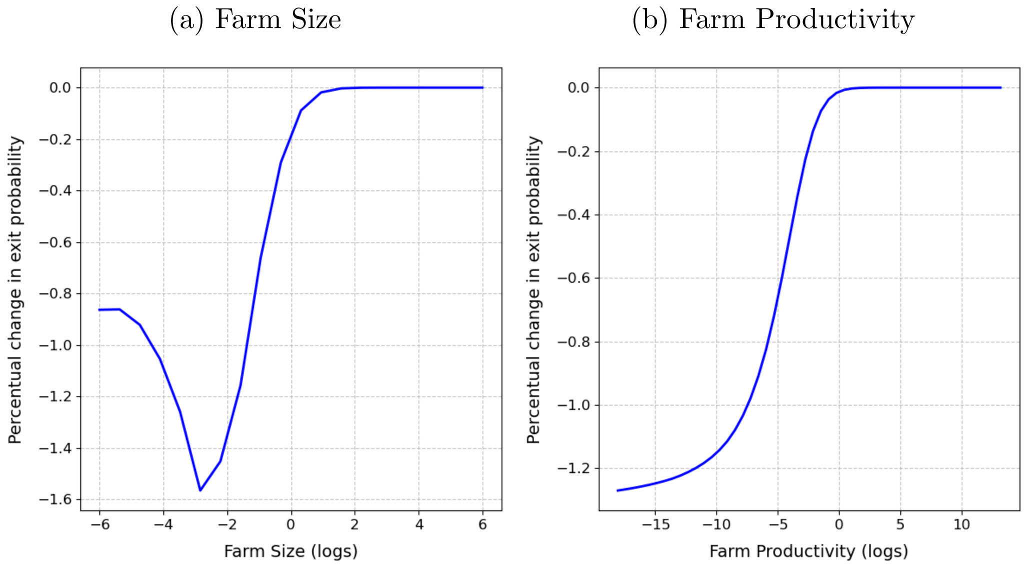
We notice that, in the data used in our reduced-form section, the average climate shock is 100 days. According to results from Table 4, an average shock leads to a 11.8 percent reduction in consumption. We therefore interpret our 25 percent reduction in aggregate agricultural productivity being roughly twice as large as the average shock estimated in the reduced-form section.

We now turn to the dynamic implications of a climate shock to the average farm size.

[22] Specifically, the amount of land that a farmer with productivity s would operate would be given by $\ell(s) = \frac{s}{\int sdG\,(\omega)} L^S$

[23] To solve for the transition dynamics, we use the Sequence Space Jacobian approach proposed in Auclert et al. (2021).

Figure 5: Impact of a Weather Shock on the Exit Probability at $t = 0$



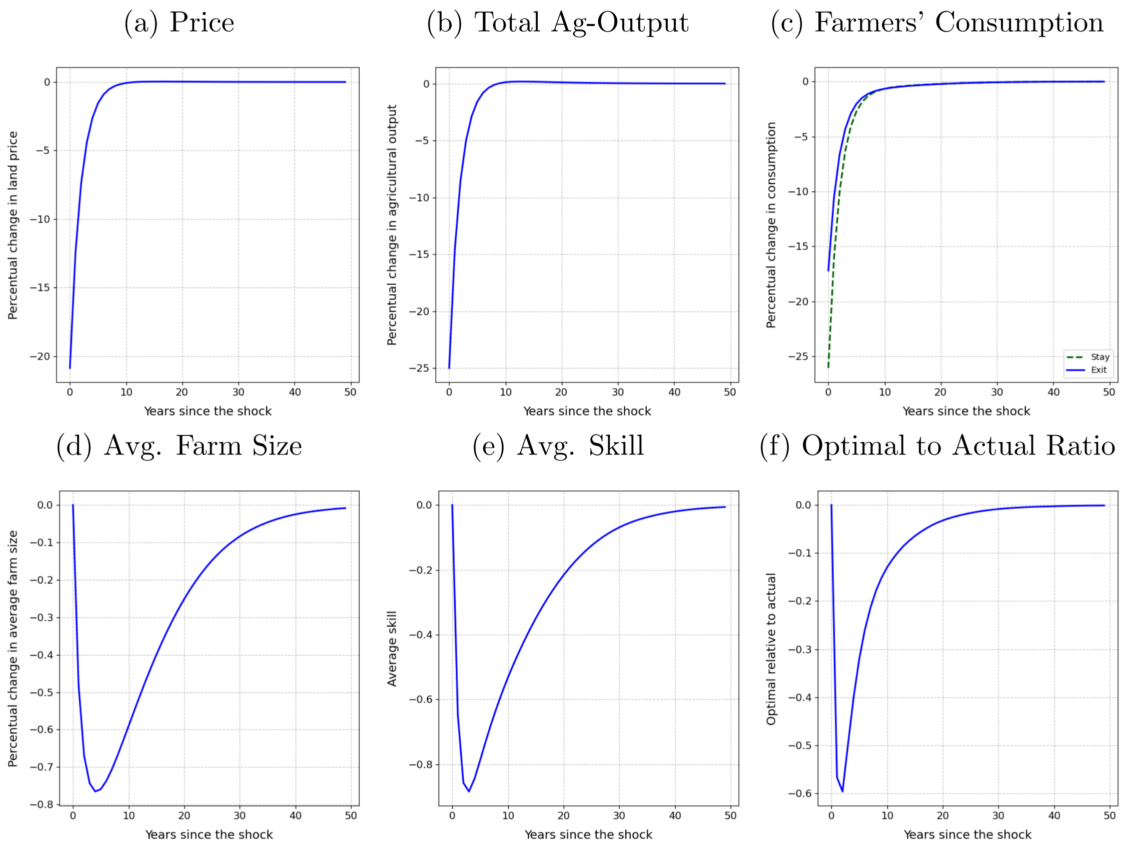
Notes: This figure shows the impact of a temporary weather shock on the change in the probability of becoming a farmer μ_t at $t = 0$. Panel (a) averages over the density of agents in terms of their landholdings. Panel (b) averages over the density of agents in terms of their farm productivity.

Figure 6 Panel (d) shows that average farm size falls by 1.2 percent. Recall that, in the reduced-form estimate, an average shock reduces average farm size by 1.6 percent—our result is therefore in the same order of magnitude as the one implied by the data. Importantly, our model implies a very strong persistence: 10 years after the shock, the average farm size is still 0.6 percent smaller.

The reason why we observe this influx of agents from urban to agriculture is primarily due to the short-term nature of the shock. When productivity falls, some farmers sell their landholdings and some of them exit agriculture altogether. That leads to a reduction in the price of land. For landless households, the value of becoming a farmer, however, is only mildly affected, because future agricultural productivity returns to its level shortly after the shock. As a consequence, this reduction in the price of land, without an equivalent drop in the present value of becoming a farmer, attracts landless households to farming. In contrast to the short-term shock, when we simulate a long-term one—i.e., when we simulate a shock with a strong a strong persistence—, we then find an net outflow of farmers from agriculture and increase in average farm size instead, since price reductions do not attract households from non-agricultural activities as much—see Appendix Figure O.4.

To better understand mechanisms, it is useful to separate the influence of a drop in the price of land p_t from the influence of a drop in aggregate productivity Z_t. Appendix OD derive analytical results that base our discussion here. Let us first focus on a short-term

Figure 6: Dynamic Effects of a Weather Shock



reduction in aggregate productivity Z_t, ignoring general equilibrium effects on the future value of becoming a farmer versus a worker or on land prices p_t. In that case, the impact on the incentives to become a farmer are unambiguous: current consumption falls, which makes farmers value more the additional consumption they obtain by choosing to exit—a result of the decreasing utility of consumption. One can think of this as agents becoming closer to subsistence level of consumption, which makes them value more the additional current consumption they get by their land to become a worker.

When we look at the impact of a reduction in the price of land p_t, ignoring changes in aggregate productivity Z_t or future value of farming and working, the impact on the probability of becoming a farmer is ambiguous. On the one hand, agents become poorer, since their asset drops in value and they are unable to consume as much by selling their land—similarly to the reduction in productivity Z_t, that mechanism comes from the decreasing utility of consumption. On the other hand, because land prices are lower, the purchasing power of agents rise, which increases their incentives to become a farmer. Notice that workers are only influenced by this last mechanism since they do not own any land. As such a reduction in the price of land, holding everything else constant, will unambiguously increase their probability of becoming a farmer.

Lastly, we turn to the determinants of aggregate agricultural productivity. The average skill of agents drops by 0.7 percent in the aftermath of the climate shock. That occurs because agents in the urban sector, who have a comparative advantage in the urban sector relative to agriculture, enter in the agricultural economy as a response to the shock, but they take long to exit. As a result, even as long as 20 years after the shock, the average skill of the farmers are smaller.

When we look at optimality *within* the agricultural sector, which considers the dispersion in MPL *given* the pool of farm productivity in the economy, we observe an improvement. Further inspection shows that this occurs because the negative shocks disproportionately remove agents with low farm productivity, who have a high MPL — see Appendix Figure 5(b). The new farmers who enter the economy have in average lower farm productivity, but they also own smaller farms.

7 Conclusion

This paper explores the effect of uninsured weather shocks on distress sales and farm size dynamics. 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 households. Consistent with the aggregate patterns we find on land sales and land distribution, we also show that these shocks induce rural households to migrate, and decrease household's consumption and asset ownership.

Distress sales after a negative productivity shock might depress land prices. Our results show that land *fragmentation* takes place after such shocks. We rationalize these results with dynamic, heterogeneous household model in which uninsured farmers make landholding and occupational choices. The combination of occupational choices with sector-specific shocks leads to an expansion of the aggregate supply of land, a temporary drop in land values, and to a net increase in the number of households occupied in agriculture operating relatively smaller farms.

Our model results suggest 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.

References

Acampora, M., Casaburi, L., and Willis, J. (2022). Land rental markets: Experimental evidence from kenya. Technical report, National Bureau of Economic Research.

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

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.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–684.

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, A. (2017). Fondo de Tierras del Acuerdo Agrario de La Habana: Estimaciones y propuestas alternativas. *Documento CEDE*, No. 41.

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2021). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. *Econometrica*, 89(5):2375–2408.

Avila, A. F. D. and Evenson, R. E. (2010). Total factor productivity growth in agriculture: The role of technological capital. *Handbook of agricultural economics*, 4:3769–3822.

Bolhuis, M. A., Rachapalli, S. R., and Restuccia, D. (2021). Misallocation in indian agriculture. Technical report, National Bureau of Economic Research.

Brooks, W. and Donovan, K. (2020). Eliminating uncertainty in market access: The impact of new bridges in rural nicaragua. *Econometrica*, 88(5):1965–1997.

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. 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.

Clementi, G. L. and Palazzo, B. (2016). Entry, exit, firm dynamics, and aggregate fluctuations. *American Economic Journal: Macroeconomics*, 8(3):1–41.

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

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

DANE (2019). Encuesta nacional agropecuaria, reporte tecnico. Technical report, Departamento Nacional de Estadística, Colombia.

de Roux, N. (2021). Exogenous shocks, credit reports and access to credit: Evidence from colombian coffee producers. *Documento CEDE*, (57).

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.

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.

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

Greenwood, J., Kircher, P., Santos, C., and Tertilt, M. (2019). An equilibrium model of the african hiv/aids epidemic. *Econometrica*, 87(4):1081–1113.

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.*

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and structural transformation. *Handbook of economic growth*, 2:855–941.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150.

Ibánez, A., Munoz-Mora, J., and Gafaro, 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.

Jaimovich, N., Terry, S. J., and Vincent, N. (2023). The empirical distribution of firm dynamics and its macro implications. Technical report, National Bureau of Economic Research.

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.

Kaboski, J. P., Lipscomb, M., Midrigan, V., and Pelnik, C. (2022). How important are investment indivisibilities for development? experimental evidence from uganda. Technical report, National Bureau of Economic Research.

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.

Khan, A. and Thomas, J. K. (2008). Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2):395–436.

Krishna, A. (2010). *One illness away: Why people become poor and how they escape poverty.* OUP Oxford.

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*.

Manysheva, K. (2022). Land property rights, financial frictions, and resource allocation in developing countries. Technical report, Working Paper.

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

Mazur, K. and Tetenyi, L. (2024). The macroeconomic impact of agricultural input subsidies. *Working Paper*.

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.

Pellegrina, H. S. (2022). Trade, productivity, and the spatial organization of agriculture: Evidence from brazil. *Journal of Development Economics*, 156:102816.

Peralta-Alva, A., Tam, X. S., Tang, X., and Tavares, M. M. (2023). Tax revenues in low-income countries. *The Economic Journal*, 133(653):2001–2024.

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.

Online Appendices for “Farm Size Distribution, Weather Shocks, and Agricultural Productivity”

Not for Publication

Julián Arteaga, Nicolás de Roux, Margarita Gáfaro,
Ana María Ibáñez and Heitor S. Pellegrina *

Table of Contents

OA Data **3**
OA.1 Land Transaction Data. 3
OA.2 Farm Size Distribution. 4
OA.3 Land Transaction Data - Landless Buyers. 5
OA.4 Weather Shocks. 6
OA.5 Farm-Level. 7

OB Additional Reduced-Form Results **7**
OB.1 Flexible specification 7
OB.2 Robustness in Farm Size Distribution and Land Transaction 9
OB.3 Robustness in New Entry 9
OB.4 The Land Ceiling Regulation 10

OC Estimating ρ **10**

OD Analytical Results **11**
OD.1 TFP shock and the probability of becoming a farmer. 11
OD.2 Price shock and the probability of becoming a farmer. 12

OE Tables and Figures **13**

List of Tables

O.1 Effects of Weather Shocks on Farmers’ Exit and Farm Size Growth 13
O.2 Descriptive Statistics 14
O.3 Effect of Temperature Shock on Land sales: matching vs. non-matching buyers 15
O.4 Temperature Shocks and Land Sales - Alternative Definitions of Shocks 16
O.5 Temperature Shocks and Average Farm Size - Alternative Definitions of Shocks 17

*Email: jgarteaga@worldbank.org, nicolas.de.roux@uniandes.edu.co, mgafargo@banrep.gov.co, anaib@iadb.org, hpelleg3@nd.edu.

O.6 Temperature Shocks and Land Sales - Alternative Specifications 18
O.7 Temperature Shocks and Average Farm Size - Alternative Specifications 19
O.8 Temperature Shocks and Household Assets 20
O.9 Temperature Shocks, Farm Size, and Share of Government-Allocated Area . . . 20
O.10 Temperature Shocks and Farm Size - Heterogeneous Effects by Contiguous Plots 21
O.11 Temperature Shocks and Number of owners, by Initial Size Quantile 21

List of Figures

O.1 Effects of Weather Shocks on Probability of being a Farmer 22
O.2 One Century of Land Allocations - 1901–2012 22
O.3 Ratio of Land Sales to Number of Allocations 23
O.4 Agricultural Productivity Shock with Strong Persistence 24
O.5 Mean and Median Allocation Size - 1961–2012 25
O.6 Number of Allocations (SIDER) vs. Number of Registrations (SNR) 25
O.7 Temperature Shocks Across Space - 2000 and 2010 26
O.8 Temperature Shocks Across Time 27
O.9 Temperature Shocks and Land Transactions 27
O.10 Temperature Shocks and Farm Size 28
O.11 Temperature Shocks and Number of Owners by Initial Distribution Quantiles - Alternative Partitions 28

OA Data

This section describes in detail all the sources of data that we use from Colombia. Appendix Table O.2 offers summary statistics for our main datasets.

OA.1 Land Transaction Data.

Recipients of land in Colombia must register the property in the office of the local public notary, and all formal land transactions carried out over the estate, including mortgages, must be registered in the national land registry mantained by the SNR. Our land transaction dataset contains the whole transaction history of all land plots granted by the government to private individuals at any point between 1900 and 2010. We match the location of the property in the SNR dataset to the official list of Colombian municipalities provided by DANE, Colombia's National Statistical Agency.[24] Using this information, we construct a balanced yearly panel both at the municipality with information on the number of full and partial land sales, mortgages, and government land allocations.

The Colombian government has carried out the free allocation of public idle lands (*baldíos*) to private individuals uninterruptedly since the beginning of the twentieth century. These 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 O.2 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 properties have been granted to private individuals in 1,034 of the 1,122 existing municipalities. These properties 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).

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

[24]Municipalities are the smallest official administrative division in Colombia. There are approximately 30,000 veredas in Colombia and 1,123 municipalities.

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 land 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 O.5 shows the evolution of the average and median size of allocated properties since 1960. The overwhelming majority of land allocations made throughout 1961–2014 period consisted of relatively small land properties, 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 OB.4, 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 property, beneficiaries must register the property in the office of the local public notary, and all formal land transactions carried out over the property (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.[25]

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.[26] 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 properties. 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 property's location, the date in which it occurred, and the type of transaction.

OA.2 Farm Size Distribution.

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 farm-level

[25] The history of the transactions carried out over a property, 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.

[26] 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 O.6 shows that allocations and real estate registrations follow each other closely across time, suggesting that the great majority of land properties allocated did end up being registered.

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.[27] 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 farms 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 farms, the number of owners and average farm 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$^{\text{th}}$ percentile) properties registered. Our final sample is made up of 927 municipalities, which encompass 85.3% of the rural population in the country.

OA.3 Land Transaction Data - Landless Buyers.

The observed increase in the number of landowners per municipality shown in Figure O.10 and Table 2 indicates that large landowners *within* the municipality are not expanding their operations by purchasing the small plots sold after the shocks. We cannot rule out, however, the possibility that these plots are being acquired by large landowners from nearby municipalities.

To investigate this, we compile yearly lists of landowners at the *departamento* level using the history of land transactions. We define an individual as a land owner if it appears in the land registry data at the receiving end of a transaction—be it a sale, an inheritance or a government allocation—and build a list of current owners for every year in our sample. We then match the names of individuals buying land on a given year with the list of current landowners to determine whether such buyers owned any land in any other municipality of the departamento.

Since we do not have information on the ID numbers of buyers and owners, we match individuals in both lists by first and last name, and work under the assumption that all registries under the same name within a departamento belong to the same individual. Because names are subject to mis-spellings or changes in the order in which first and last names are recorded, we carry out this matching process

[27] 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.

using different types of string matching algorithms which vary in how conservative they are regarding the occurence of type-I errors (exact matching, bigram, and Jaro-Winkler distance) and compare the results.

OA.4 Weather Shocks.

We define temperature shocks that are specific to each administrative unit 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 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 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.[28] 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 municipality-day pair (m, d) by taking a weighted average of the pixels in the municipality using as weights the area of the pixel relative to the total area of the municipality. We compute the historical quarterly distribution of daily temperatures by considering all temperature measurements for pairs m, d in calendar-quarter q throughout the period 1979–2016. For each municipality 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 municipality-day as atypically high if it is above the 80$^{\text{th}}$ percentile of the corresponding distribution of average daily temperatures of m, 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 t, 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 municipality-year (m, t) 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. $t-1$, $t-2$). Figures O.7 and O.8 in the appendix show the spatial and temporal variation of the resulting temperature shock measures across municipalities. 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 municipality and takes

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

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 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 municipality to obtain monthly average rainfall. Again, we use as weights the size of the pixel relative to the size of the municipality. For a given year, we add across months to obtain a measure of total precipitation in the pair municipality-year (m, y).

OA.5 Farm-Level.

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-Growing, and South. Within each region, municipalities and veredas were randomly chosen. The baseline sample includes 17 municipalities. 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 inhow migration, farm size, land ownership, and household consumption change in response to temperature shocks. Panel B of O.2 contains descriptive statistics of the ELCA panel. On average, 12% of households migrated, 88% had any land and the average size of the farm was 2.5 hectares, with 75% of farms being smaller than 3 hectares.

OB Additional Reduced-Form Results

OB.1 Flexible specification

We estimate a flexible-lag specification that further exploits the spatial and temporal variation in the occurrence of extreme temperature events, and estimates independent coefficients for the impact of weather shocks occurring each year between t and $t-7$. This specification allows for the possibility that households may exhaust alternative coping mechanisms after consecutive years of adverse weather conditions.

Specifically we estimate a regression of the form:

$$s_{i,t} = \sum_{j=0}^{7} \beta_j \sum_{d=1}^{365} AtypicalDay_{i,d,t-j} + X'_{i,t}\delta + \eta_i + \kappa_t + \varepsilon_{i,t}, \tag{O.1}$$

where $s_{i,t}$ represents the number of land transactions (sales or mortgages), log number of farm owners, and log average and median farm size in municipality i and year t.[29] Our main variables of interest, $AtypicalDay_{i,t-j}$, represent the number of days with atypical temperatures in the municipality during year $t-j$. We include the contemporaneous value of this variable and seven lags.

The vector $X_{i,t}$ includes time-varying municipality characteristics. In particular, rainfall at t and seven lags, the cumulative number of farms allocated in the municipality up to t, which controls for the land on which we observe transactions, a dummy variable for cadastral updates, and the total municipal land area recorded in the registry, which accounts for changes in registry coverage. The model also includes municipality fixed effects η_i to control for time-invariant heterogeneity across municipalities, and year fixed effects κ_t to control for time-specific shocks to land markets and farm size common to all municipalities.

The coefficients β_j capture the reduced form effects of contemporaneous and lagged days of extreme temperature on land transactions and farm size. Although distress land sales are likely to be an immediate response to negative income shocks, notarizing a transaction and updating the property information on the land registry are the final steps in the process of buying land, which can take several months or even years.[30] We complement the results from this dynamic specification with estimates of the aggregate effect of temperature shocks over a two-year period. In particular, we estimate:

Figure O.9 presents the coefficient estimates from equation (O.1) using the total number of land sales and land mortgages per municipality as dependent variables. The figure shows that increases in the frequency of extreme temperature events lead to a rise in the number of land sales for up to two years after their occurrence. Also, extreme temperature events increase the number of land mortgages, with statistically significant effects lasting up to five years.

Figure O.10 shows the coefficient estimates of atypical temperature and its lags, with the number of landowners (panel a), and mean and median farm size (panel b) as dependent variables in equation (O.1). The figure indicates that atypical temperature days increase the number of landowners and reduce farm size, with statistically significant effects starting two years later and lasting up to seven years. The longer lag in the effects with the registry data aligns with land transactions being reported to the land registry as the final step in the purchase process. Also, since the land registry reflects the stock of properties, the shocks are likely to have permanent effects, unless further land transactions occur.

[29] We define *farms* as a piece of land with a distinct registry number.

[30] In this process, buyers and sellers first sign a transaction agreement, outlining a sequence of payment installments. Usually, signing the public deed at a notary's office takes place with the final payment. The signed deed is then submitted to the local land registry office to formalize the transfer of ownership. We observe land transactions on the date in which the public deed is signed and observe changes in average farm size and in the number of owners for the year in which the deeds are submitted to the registry office.

OB.2 Robustness in Farm Size Distribution and Land Transaction

Tables O.4 and O.5 report estimates from equation (4) using alternative definitions of temperature shocks. Panels A and B define atypical temperature using as thresholds the 5th and 95th percentiles and the mean $\pm$ 1.5 standard deviations of the temperature distribution, respectively. Panel C uses the same threshold as the main specification, but adjusts the time reference for computing the temperature distributions to 1990-2011. Panels D and E define shocks using the SPEI index.

Tables O.6 and O.7 report estimates from alternative specifications to the one in equation (4). Panel A incorporates *departamento*-specific time trends, which allows us to account for potential spurious correlations between regional temperature shock trends and our variables of interest. Panels B and C include additional controls for forced displacement and homicide rates to capture the potentially confounding relationship between violent conflict and weather shocks within municipalities. Panel D presents results from the main specification, with two-way clustered standard errors at both the municipality, and the departamento $\times$ year levels.

OB.3 Robustness in New Entry

We show here that results shown in Table 3—distinguishing which purchases are made by landless individuals and which are made by those already owning some land elsewhere—are robust to the choice of the matching algorithm and to the special sampling properties of the land transaction data.

First, given that the matching of buyers and owners is made based on first and last names which are subject to misspellings, we show results obtained from different string matching algorithms (exact match, bigram, and Jaro-Winkler) which are more or less conservative regarding the possibility of wrongly matching two distinct individuals with similar names (type II error) or failing to match the same individual across lists with a misspelled name (type I error). Panels A, B, and C of Table O.3 show that our results do not depend on the type of matching algorithm chosen.

Second, given that the land transaction data is only available for plots which were at some point allocated by the government to private individuals, it might be possible that individuals who own land but are not observed in our owner lists are incorrectly classified as landless. This omission could potentially bias the coefficient for sales to landless purchases (column 2 in Table 3) upwards.

To gauge the severity of this bias, we re-estimate these regressions using only municipalities where the share of private farmland which was part of a government allocation is above the median. We expect that in these municipalities the rate of individuals incorrectly classified as landless should be lower, given that a larger share of land transactions are observed. Columns 4, 5, and 6 in Table O.3 show the coefficients obtained from estimations carried out in this subsample alone. These results show that the share of total weather-driven land purchases made specifically by landless buyers is almost identical in both samples, suggesting any bias from this type of incorrect owner classification is low.

OB.4 The Land Ceiling Regulation

We investigate if our results on the absence of land consolidation are due to institutional factors stemming from Colombia's land regulation policies. As discussed in section **??**, 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 farms whose provenance was a government allocation.[31]

To test if these restrictions are in fact explaining our results, we re-estimate the model in (4) 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 farms. Hence, if restrictions are driving the land-fragmentation results shown in Table 2 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 O.9 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.

OC Estimating ρ

To estimate the persistence in the farm productivity shock, we use our longitudinal farm-level data, which contains data on total agricultural output per farm and the landholdings. Using that information, we construct the following measure of s_{it}

$$\bar{s}_t = Z_t \left[\frac{\bar{y}_t}{\left(\bar{\ell}_t^{\alpha} k_t^{1-\alpha} \right)^{\gamma}} \right]^{\frac{1}{1-\gamma}},$$

where $\bar{y}_t$ and $\bar{\ell}_t$ are our measured values. In what follows, we assume that Z_t and k_t are not systematically correlated with $\bar{\ell}_t$. We then postulate that the farm productivity of a farmer i in t is given by

$$\log \bar{s}_{i,t} = \rho \log \bar{s}_{i,t-1} + \sigma \epsilon_{i,t},$$

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

where $\epsilon_{i,t}$ is the productivity shock. Since the gap between years in the data is three, what we effectively observe is

$$\log \bar{s}_{i,t} = \rho^3 \left(\log \bar{s}_{i,t-3}\right) + \sigma \left(\epsilon_{i,t-2} + \epsilon_{i,t-1} + \epsilon_{i,t}\right).$$

We estimate the regression above by adding a control for the probability of exiting during these three years, which estimate using a logit model and using a polynomial of the initial farm size as the explanatory variable. Our final estimation is

$$\log \bar{s}_{i,t} = \rho^3 \left(\log \bar{s}_{i,t-3}\right) + \beta P\left(\widehat{exit_{i,t}}_{-3}\right) + \sigma \left(\epsilon_{i,t-2} + \epsilon_{i,t-1} + \epsilon_{i,t}\right).$$

where $P\left(\widehat{exit_{i,t}}_{-3}\right)$ is the predicted probability that the farmer left over the three years period.[32] This procedure gives us the value of ρ.

OD Analytical Results

To better understand how the change in price shapes agents' decision to become a farmer, we study analytically the impact of a small change in land price and the aggregate productivity. We do so based on equation (14), ignoring the general equilibrium effects, related to changes in the distribution, and using envelope conditions.

OD.1 TFP shock and the probability of becoming a farmer.

For a given household, $\omega_t = (w_t , \ell_t , s_t)$, the impact of a small change in Z_t on her probability of becoming a farmer μ_t is

$$\frac{\partial \mu_t}{\partial Z_t} = \frac{1}{\kappa} \mu_t \left(1 - \mu_t\right) \left[u'\left(c_t^F\right) - u'\left(c_t^W\right)\right] \frac{y_t}{Z_t} > 0$$

where c_t^F is the consumption if the household becomes a farmer, c_t^W is the consumption if the household becomes a worker, and $u'()$ is the marginal utility of consumption. Notice that, because $y_t = 0$ for workers, they are unaffected by the shock.

In the equation above, we notice that $u'\left(c^F\left(w_t, \ell_t, s_t\right)\right) > u'\left(c^W\left(w_t, \ell_t, s_t\right)\right)$ since consumption is always lower if households buy land. Therefore, the impact of a reduction in Z_t is to always reduce the probability that agents will become a farmer.

Notice that the magnitude of the impact of Z_t on $\mu\left(w_t, \ell_t, s_t\right)$. On the one hand, more productive farmers are more severely affected by the shock, since they have larger s_t. On the other hand, the magnitude of the difference between $u'\left(c^F\left(w_t, \ell_t, s_t\right)\right) - u'\left(c^W\left(w_t, \ell_t, s_t\right)\right)$ is larger for less productive farms, because of the curvature of the utility function.

[32] Controlling for the predicted probability decreases the correlation between shocks, but the change in point-estimates are, in practice, small.

OD.2 Price shock and the probability of becoming a farmer.

We now turn to the impact of a small shock to land prices on the probability of becoming a farmer

$$\frac{\partial \mu_t}{\partial p_t} = \frac{1}{\kappa}\mu_t\left(1-\mu_t\right)\left[\underbrace{\left[u'\left(c_t^F\right)-u'\left(c_t^W\right)\right]\ell_t}_{\text{income effect (+)}} - \underbrace{u'\left(c^F\right)\ell_t^*}_{\text{purchasing power effect (-)}}\right] \lesseqgtr 0.$$

Now, we have two different mechanism. First, we have an income effect, which comes the fact that agents become richer when they own more land. This mechanism moves in the same direction of prices and its influence is larger when farmers are poorer, since the difference between in marginal utility of consumption $u'\left(c^F\right)-u'\left(c^W\right)$ is larger. Notice that agents with no land are not influenced by the income effect.

Second, we have a purchasing power effect, which moves in the opposite direction of prices. When prices drop, agents are more likely to stay, since they can purchase more land. Notice that, for agents outside of agriculture, this is the only mechanism that influences their decision, since they own no wealth from land.

OE Tables and Figures

Table O.1: Effects of Weather Shocks on Farmers' Exit and Farm Size Growth

	Exit probability		Log farm size	
	(1) Colombia	(2) West Africa	(3) Colombia	(4) West Africa
$TempShocks$	0.192***	0.096***	-0.292	-0.110
	(0.057)	(0.013)	(0.237)	(0.068)
$Tempshocks \times LF$	-0.037***	-0.027***	0.731***	0.509***
	(0.014)	(0.002)	(0.085)	(0.016)
Observations	7,336	50,742	6,782	43,640
R^2	.55	.55	.83	.74
Mean Shock Var	2.612	3.297	2.606	3.284
SD Shock Var	0.20	0.39	0.19	0.39

Notes: Estimates of equation (1) on the probability of exiting from agriculture (defined as a household reporting no landholdings) and farmers' log land size. LF_f is an indicator variable for whether the farm is above the average farm in its respective country. *Mean Shock Var.*, and *SD Shock Var.* show the average and standard deviation in the number of days (measured in hundreds) with atypical temperature in each sample. Standard errors clustered at the geographic administrative level in which the weather shock is measured in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table O.2: Descriptive Statistics

	Mean (1)	Std. Dev. (2)	Min (3)	Max (4)
	Panel A: Municipality (N= 864)			
SNR				
Number of Total Land Sales	12.62	24.75	0	292
Number of Full Land Sales	10.83	21.63	0	281
Number of Partial Land Sales	1.79	6.05	0	133
Number of Mortages	2.62	7.54	0	172
Land Registry				
Number of owners	2,501.13	2,160.43	21	18,768
Mean of farm size	32.06	106.71	1	1,693
Median of farm size	16.15	86.73	0	1,438
Controls				
Number of total allocations	444.68	679.23	1	6,550
=1 if land registry update	0.07	0.25	0	1
Registered area (1000 ha.)	41,255.76	87,792.98	340	1,475,761
Accumulated precipitation	3,516.60	2,647.67	372	21,144
Days of atypical high temperature	120.47	75.50	5	483
Days of atypical low temperature	154.92	72.23	1	401
Days of atypical temperature	275.39	43.56	157	485
	Panel B: ELCA - Household (N= 4,293)			
=1 if HH has land	0.88	0.33	0	1
=1 if farm size $\leq$ 3 ha	0.75	0.43	0	1
=1 if HH migrated	0.12	0.33	0	1
=1 if HH sold animals	0.69	0.46	0	1
Asset index	0.00	0.39	-1	3
Farm size (ha.)	2.52	5.53	0	118
Consumption per capita	2.66	2.10	0	54
Accumulated precipitation	3,746.58	2,580.70	720	22,934
Days of atypical high temperature	267.29	119.26	64	528
Days of atypical low temperature	49.76	43.33	0	176
Days of atypical temperature	317.04	85.59	178	529

Notes: Summary statistics for each estimation sample. Panel A describes the variables used for municipality-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 farms in the municipality from 1901 until the year of observation. Panel B 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 ($t-2$, $t-1$) with either abnormally high or low temperatures. Accumulated precipitation is the volume of rain in milliliters for year t.

Table O.3: Effect of Temperature Shock on Land sales: matching vs. non-matching buyers

	Full sample			Above median land share from govt. allocations		
	(1)	(2)	(3)	(4)	(5)	(6)
	Total Sales	Sales to Landless Buyers	Sales to Already Owners	Total Sales	Sales to Landless Buyers	Sales to Already Owners
Panel A: Fuzzy Matching - Bigram						
$TempShock_{i,t}$	2.537***	1.866***	0.671***	3.524***	2.679***	0.845***
	(0.536)	(0.408)	(0.192)	(0.855)	(0.654)	(0.300)
Observations	10,021	10,021	10,021	5,097	5,097	5,097
R^2	0.905	0.889	0.838	0.899	0.877	0.839
Mean Dep. Var	12.62	9.37	3.25	20.91	15.47	5.44
Panel B: Fuzzy Matching - Jaro Winkler						
$TempShock_{i,t}$	2.537***	1.402***	1.135***	3.524***	2.080***	1.445***
	(0.536)	(0.325)	(0.291)	(0.855)	(0.526)	(0.466)
Observations	10,021	10,021	10,021	5,097	5,097	5,097
R^2	0.905	0.857	0.889	0.899	0.845	0.885
Mean Dep. Var	12.62	6.10	6.52	20.91	10.04	10.87
Panel C: Exact Matching						
$TempShock_{i,t}$	2.537***	1.780***	0.757***	3.524***	2.546***	0.978***
	(0.536)	(0.412)	(0.190)	(0.855)	(0.663)	(0.297)
Observations	10,021	10,021	10,021	5,097	5,097	5,097
R^2	0.905	0.889	0.837	0.899	0.877	0.839
Mean Dep. Var	12.62	9.62	3.00	20.91	15.90	5.01

Notes: Data from the National Superintendency of Notaries (SNR) records. Coeficient estimates from equation (4). The main independent variable is the total number of atypical temperature days in the past two years ($t-1$, $t-2$) divided by 100. Controls are accumulated allocations, area covered by the land registry, an indicator of registry updates, accumulated precipitation and five lags. See texts for more details. Standard errors clustered at the municipality level reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Table O.4: Temperature Shocks and Land Sales - Alternative Definitions of Shocks

	Total (1)	Full (2)	Partial (3)	Mortg. (4)
Panel A: Threshold 5, 95				
$TempShocks$	4.290***	3.150***	1.140***	1.922***
	(0.867)	(0.804)	(0.370)	(0.368)
Observations	10,021	10,021	10,021	10,021
R^2	0.906	0.902	0.636	0.757
Mean Dep. Var	12.62	10.83	1.79	2.62
Panel B: Threshold 1.5SD				
$TempShocks$	3.689***	2.765***	0.925***	1.762***
	(0.752)	(0.695)	(0.317)	(0.314)
Observations	10,021	10,021	10,021	10,021
R^2	0.906	0.902	0.636	0.758
Mean Dep. Var	12.62	10.83	1.79	2.62
Panel C: Shorter time window				
$TempShocks$	2.248***	1.887***	0.361	1.004***
	(0.546)	(0.504)	(0.238)	(0.231)
Observations	10,021	10,021	10,021	10,021
R^2	0.905	0.902	0.635	0.757
Mean Dep. Var	12.62	10.83	1.79	2.62
Panel D: SPEI continuous				
$SPEI(-)$	0.859**	0.441	0.418**	-0.349*
	(0.349)	(0.324)	(0.171)	(0.183)
Observations	10,021	10,021	10,021	10,021
R^2	0.905	0.902	0.636	0.756
Mean Dep. Var	12.62	10.83	1.79	2.62
Panel E: SPEI No discrete				
Months($SPEI < -2$)	0.846**	0.647**	0.199	0.389***
	(0.348)	(0.322)	(0.160)	(0.121)
Observations	10,021	10,021	10,021	10,021
R^2	0.905	0.902	0.635	0.756
Mean Dep. Var	12.62	10.83	1.79	2.62

Notes: Data from the National Superintendency of Notaries (SNR) records. Coefficient estimates from equation 4. Each panels use a different temperature shock definitions. Panels A and B compute the temperature shock with the same procedure as the main specification, except they take different thresholds to define when a day has atypically low or high temperature. In panel A, the thresholds are the 5^{th} and 95^{th} percentiles of the temperature distribution, respectively. In panel B, the thresholds are the *mean -/+ 1.5 standard deviation* of the temperature distribution. Panel C computes the temperature shock with the same procedure as the main specification, except the time reference for the shock thresholds is 1990-2011. Panel D uses a variable with the negative values of the annual SPEI in years $t-1$ and $t-2$ multiplied by -1 and zero otherwise. In panel E, defines a shock measure equal to one when the monthly SPEI < -2 in the years $t-1$ and $t-2$. All regressions include the same controls and fixed effects as the main specification. Except for panel D that includes in addition a variable with the positive side of the annual SPEI and zero otherwhise. *Mean Dep. Var.* is the mean of the dependent variable. Standard errors clustered at the municipality level reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Table O.5: Temperature Shocks and Average Farm Size - Alternative Definitions of Shocks

	(1) Number of Owners (1)	(2) Mean Farm Size (2)	(3) Median Farm Size (3)
Panel A: Threshold 5, 95			
TempShocks	0.024***	-0.028***	-0.036**
	(0.008)	(0.008)	(0.015)
Observations	10,021	10,021	10,021
R^2	0.992	0.993	0.974
Mean Dep. Var	2,501.13	32.06	16.15
Panel B: Threshold 1.5SD			
TempShocks	0.016**	-0.018***	-0.027**
	(0.006)	(0.007)	(0.014)
Observations	10,021	10,021	10,021
R^2	0.992	0.993	0.974
Mean Dep. Var	2,501.13	32.06	16.15
Panel C: Shorter time window			
TempShocks	0.013***	-0.013**	-0.021*
	(0.005)	(0.005)	(0.012)
Observations	10,021	10,021	10,021
R^2	0.992	0.993	0.974
Mean Dep. Var	2,501.13	32.06	16.15
Panel D: SPEI continuous			
$SPEI(-)$	0.013**	-0.017**	-0.042**
	(0.006)	(0.007)	(0.019)
Observations	10,021	10,021	10,021
R^2	0.992	0.993	0.975
Mean Dep. Var	2,501.13	32.06	16.15
Panel E: SPEI discrete			
Months($SPEI < -2$)	0.006*	-0.006	-0.012
	(0.003)	(0.004)	(0.009)
Observations	10,021	10,021	10,021
R^2	0.992	0.993	0.974
Mean Dep. Var	2,501.13	32.06	16.15

Notes: Data from the National Land Registry (*Catastro Nacional*), mantained by the National Geographical Institute (IGAC). Panels try different temperature shock definitions from the main specification. Panels A and B compute the temperature shock with the same procedure as the main specification, except they take different thresholds to define when a day has atypically low or high temperature. In panel A, the thresholds are the 5^{th} and 95^{th} percentiles of the temperature distribution, respectively. In panel B, the thresholds are the *mean - 1.5 standard deviation* and the *mean + 1.5 standard deviation* of the temperature distribution, respectively. Panel C computes the temperature shock with the same procedure as the main specification, except the time reference for the shock thresholds is 1990-2011. Panel D uses a variable with the negative values of the annual SPEI in years $t-1$ and $t-2$ multiplied by -1 and zero otherwise. In panel E, defines a shock measure equal to one when the monthly SPEI < -2 in the years $t-1$ and $t-2$. All regressions include the same controls and fixed effects as the main specification. Except for panel D that includes in addition a variable with the positive side of the annual SPEI and zero otherwhise. *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.

Table O.6: Temperature Shocks and Land Sales - Alternative Specifications

	Total (1)	Full (2)	Partial (3)	Mortg. (4)
Panel A: Controls for departamento specific linear trends				
TempShocks	2.532***	2.018***	0.514**	1.059***
	(0.535)	(0.502)	(0.224)	(0.241)
Observations	10,021	10,021	10,021	10,021
R^2	0.905	0.902	0.636	0.757
Mean Dep. Var	12.62	10.83	1.79	2.62
Panel B: Controls for displacement				
TempShocks	2.524***	2.000***	0.525**	1.049***
	(0.537)	(0.505)	(0.229)	(0.237)
Observations	10,014	10,014	10,014	10,014
R^2	0.906	0.902	0.636	0.757
Mean Dep. Var	12.63	10.84	1.79	2.62
Panel C: Controls for homicide				
TempShocks	2.457***	1.938***	0.519**	1.016***
	(0.540)	(0.509)	(0.226)	(0.237)
Observations	10,014	10,014	10,014	10,014
R^2	0.906	0.903	0.636	0.758
Mean Dep. Var	12.63	10.84	1.79	2.62
Panel D: Municipality and departamento × year clusters				
TempShocks	2.537***	2.013***	0.523*	1.046***
	(0.786)	(0.732)	(0.284)	(0.388)
Observations	10,021	10,021	10,021	10,021
R^2	0.905	0.902	0.636	0.757
Mean Dep. Var	12.62	10.83	1.79	2.62

Notes: Data from the National Superintendency of Notaries (SNR) records. Panels present different robustness checks to the main specification with all controls. Panel A includes the interaction between the administrative division *Departamento* and a linear trend. Panel B and C include a measure of displaced population and total number of homicides, respectively (both as proportion of the total population in 2005 in the municipality). Panel D clusters standard errors by municipality and by departamento-year. *Mean Dep. Var.* is the mean of the dependent variable. Clustered standard errors are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Table O.7: Temperature Shocks and Average Farm Size - Alternative Specifications

	Number of Owners (1)	Mean Farm Size (2)	Median Farm Size (3)
Panel A: Controls for departamento specific linear trends			
$TempShocks$	0.015***	-0.015***	-0.022*
	(0.005)	(0.005)	(0.012)
Observations	10,021	10,021	10,021
R^2	0.992	0.993	0.975
Mean Dep. Var	2,501.13	32.06	16.15
Panel B: Controls for displacement			
$TempShocks$	0.015***	-0.015***	-0.022*
	(0.005)	(0.005)	(0.012)
Observations	10,014	10,014	10,014
R^2	0.992	0.993	0.974
Mean Dep. Var	2,501.38	32.06	16.15
Panel C: Controls for homicide			
$TempShocks$	0.015***	-0.015***	0.000
	(0.005)	(0.005)	(0.002)
Observations	10,014	10,014	10,014
R^2	0.992	0.993	0.926
Mean Dep. Var	2,501.38	32.06	0.98
Panel D: Municipality and departamento $\times$ year clusters			
$TempShocks$	0.015**	-0.015**	0.000
	(0.006)	(0.007)	(0.003)
Observations	10,021	10,021	10,021
R^2	0.992	0.993	0.926
Mean Dep. Var	2,501.13	32.06	0.98

Notes: Data from the National Land Registry (*Catastro Nacional*), maintained by the National Geographical Institute (IGAC). Panels present different robustness checks to the main specification with all controls. Panel A includes the interaction between the administrative division *Departamento* and a linear trend. Panel B and C include a measure of displaced population and total number of homicides, respectively (both as proportion of the total population in 2005 in the municipality). Panel D clusters standard errors by municipality and by departamento-year. *Mean Dep. Var.* is the mean of the dependent variable. Clustered standard errors are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Table O.8: Temperature Shocks and Household Assets

	Refrigerators (1)	Washing Machines (2)	Blenders (3)	Ovens (4)	Microwave (5)	Heaters (6)	Showers (7)	Air-conditioning (8)	Televisions (9)	Radios (10)
TempShocks	-0.059***	0.061***	-0.043**	-0.014	-0.021***	-0.020***	-0.067***	-0.100***	-0.040**	0.003
	(0.018)	(0.019)	(0.021)	(0.009)	(0.007)	(0.005)	(0.011)	(0.010)	(0.018)	(0.023)
Observations	11,422	11,422	11,422	11,422	11,422	11,422	11,422	11,422	11,422	11,422
R^2	0.681	0.645	0.559	0.447	0.443	0.405	0.572	0.445	0.536	0.540
Mean Dep. Var	0.64	0.24	0.71	0.04	0.03	0.02	0.06	0.05	0.84	0.54
	Sound Equipment	Video Equipment	Computers	Bikes	Motorcycles	Automobiles	Houses	Office	Lots	Other Goods
TempShocks	-0.041*	0.005	-0.045***	-0.015	0.041**	-0.036***	-0.006	0.001	-0.006	-0.173***
	(0.021)	(0.020)	(0.012)	(0.021)	(0.019)	(0.009)	(0.011)	(0.003)	(0.006)	(0.013)
Observations	11,422	11,422	11,422	11,422	11,422	11,422	11,422	11,422	11,422	11,422
R^2	0.586	0.537	0.528	0.602	0.685	0.706	0.529	0.456	0.430	0.495
Mean Dep. Var	0.32	0.25	0.08	0.40	0.32	0.05	0.05	0.00	0.01	0.06

Notes: Data from ELCA. Dependent variables are dummies for the options given to the household when asked: *Does this household own any of the following assets?* Regressions include the same controls and fixed effects as the main ELCA specifications. *Mean Dep. Var.* is the mean of the dependent variable. Standard errors clustered at the household level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Table O.9: Temperature Shocks, Farm Size, and Share of Government-Allocated Area

	Control: Share allocates			H_i: Share allocated		
	Number of Owners (1)	Mean Farm Size (2)	Median Farm Size (3)	Number of Owners (4)	Mean Farm Size (5)	Median Farm Size (6)
$TempShocks$	0.014***	-0.015***	-0.022*	0.014***	-0.016***	-0.031***
	(0.005)	(0.005)	(0.012)	(0.005)	(0.005)	(0.011)
$TempShocks \times H_i$				0.000	0.004	0.027*
				(0.008)	(0.009)	(0.015)
Observations	10,021	10,021	10,021	10,021	10,021	10,021
R^2	0.992	0.993	0.975	0.992	0.993	0.975
Mean Dep. Var	2,501.13	32.06	16.15	2,501.13	32.06	16.15

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 ($t-1$, $t-2$) divided by 100. Controls are log accumulated allocations, log registry area, an indicator for cadastral update, and log accumulated precipitation with 5 lags. All regressions also include year and municipality fixed effects. *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.

Table O.10: Temperature Shocks and Farm Size - Heterogeneous Effects by Contiguous Plots

	Number of Owners (1)	Mean Farm Size (2)	Median Farm Size (3)
$TempShocks$	0.019***	-0.019***	-0.022
	(0.006)	(0.007)	(0.015)
$TempShocks \times High$	-0.003	0.005	0.010
	(0.006)	(0.006)	(0.013)
Observations	8,575	8,575	8,575
R^2	0.991	0.993	0.973
Mean Dep. Var	2,532.27	30.81	15.62

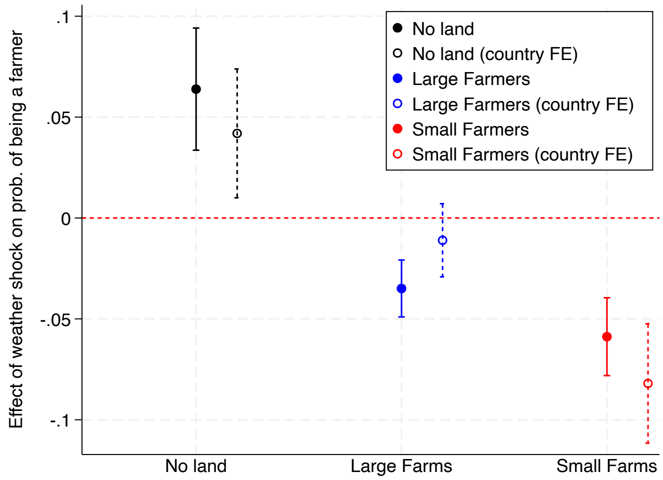
Notes: Data from the National Land Registry (*Catastro Nacional*), mantained by the National Geographical Institute (IGAC). All dependent variables are in logarithms. The main independent variable is the total number of atypical temperature days in the past two years ($t-1$, $t-2$) divided by 100. Controls are log accumulated precipitation during and five lags, log accumulated allocations, log area covered by the land registry and an indicator of registry updates. Regressions also include year and geographic fixed effects. *High* indicates a dummy variable equal to one for municipalities with a high (above the median) share of farms in the 10th percentile of the size distribution that are contiguous to at least one farm above the 90th. Neighbors are constructed from buffers around GPS coordinates of farms in the data from the National Agricultural Census of 2014. Standard errors clustered at the municipality level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Table O.11: Temperature Shocks and Number of owners, by Initial Size Quantile

	Number of owners by initial distribution quantiles (q_m^j)									
	q_m^1 (1)	q_m^2 (2)	q_m^3 (3)	q_m^4 (4)	q_m^5 (5)	q_m^6 (6)	q_m^7 (7)	q_m^8 (8)	q_m^9 (9)	q_m^{10} (10)
$TempShocks$	0.032**	0.017*	0.026***	0.016*	0.022***	0.008	0.006	-0.001	0.001	-0.000
	(0.016)	(0.010)	(0.008)	(0.008)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.004)
Observations	10,004	9,967	9,942	9,893	9,996	9,958	10,017	9,981	9,996	10,017
R^2	0.940	0.971	0.981	0.982	0.986	0.985	0.985	0.990	0.990	0.993

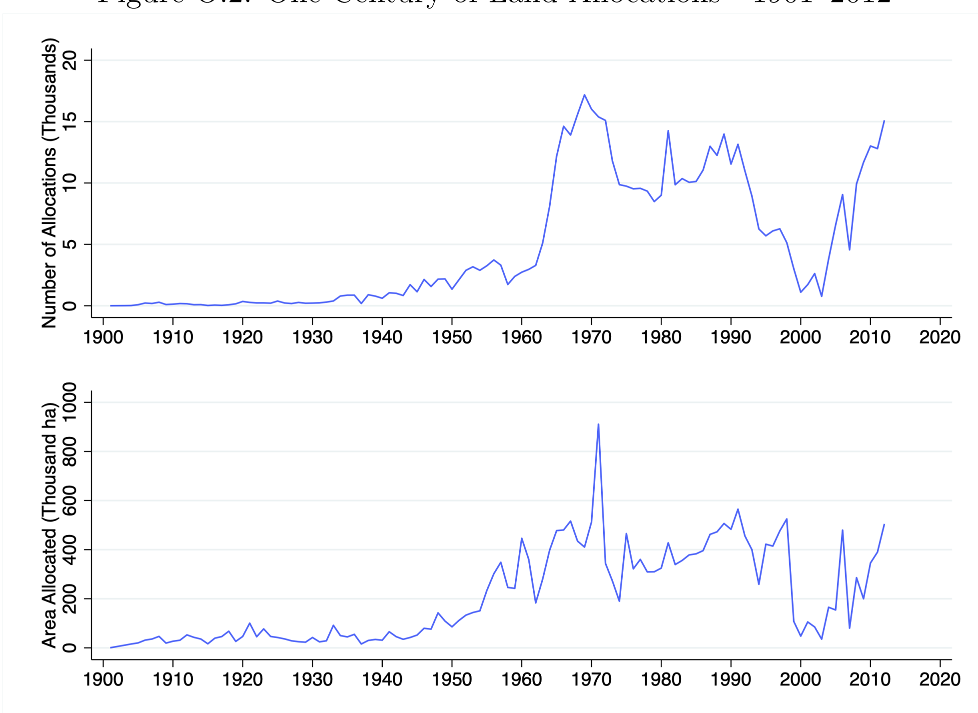
Notes: Data from the National Land Registry (*Catastro Nacional*), mantained by the National Geographical Institute (IGAC). Dependent variables are log number of owners whose farm is in the corresponding size range defined by the quantiles of the initial farm distribution. The main independent variable is the total number of atypical temperature days in the past two years ($t-1$, $t-2$) divided by 100. Controls are log accumulated allocations, log area covered by the land registry, an indicator of registry updates, log accumulated precipitation and five lags. Regressions also include year and municipality fixed effects. Standard errors clustered at the municipality level are reported in parenthesis. *p<0.1, **p<0.05, ***p<0.01.

Figure O.1: Effects of Weather Shocks on Probability of being a Farmer



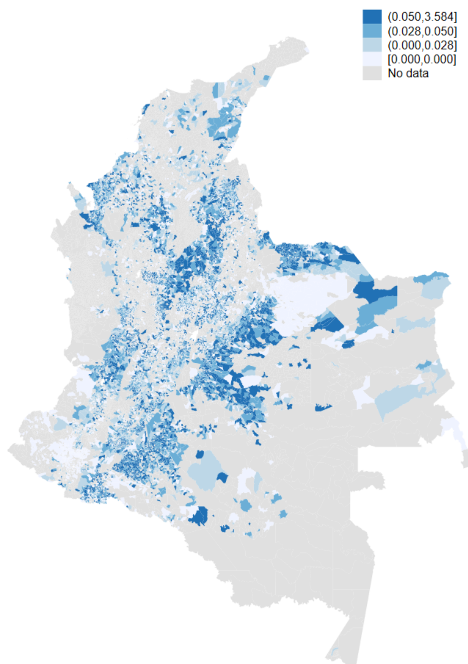
Notes:.

Figure O.2: One Century of Land Allocations - 1901–2012



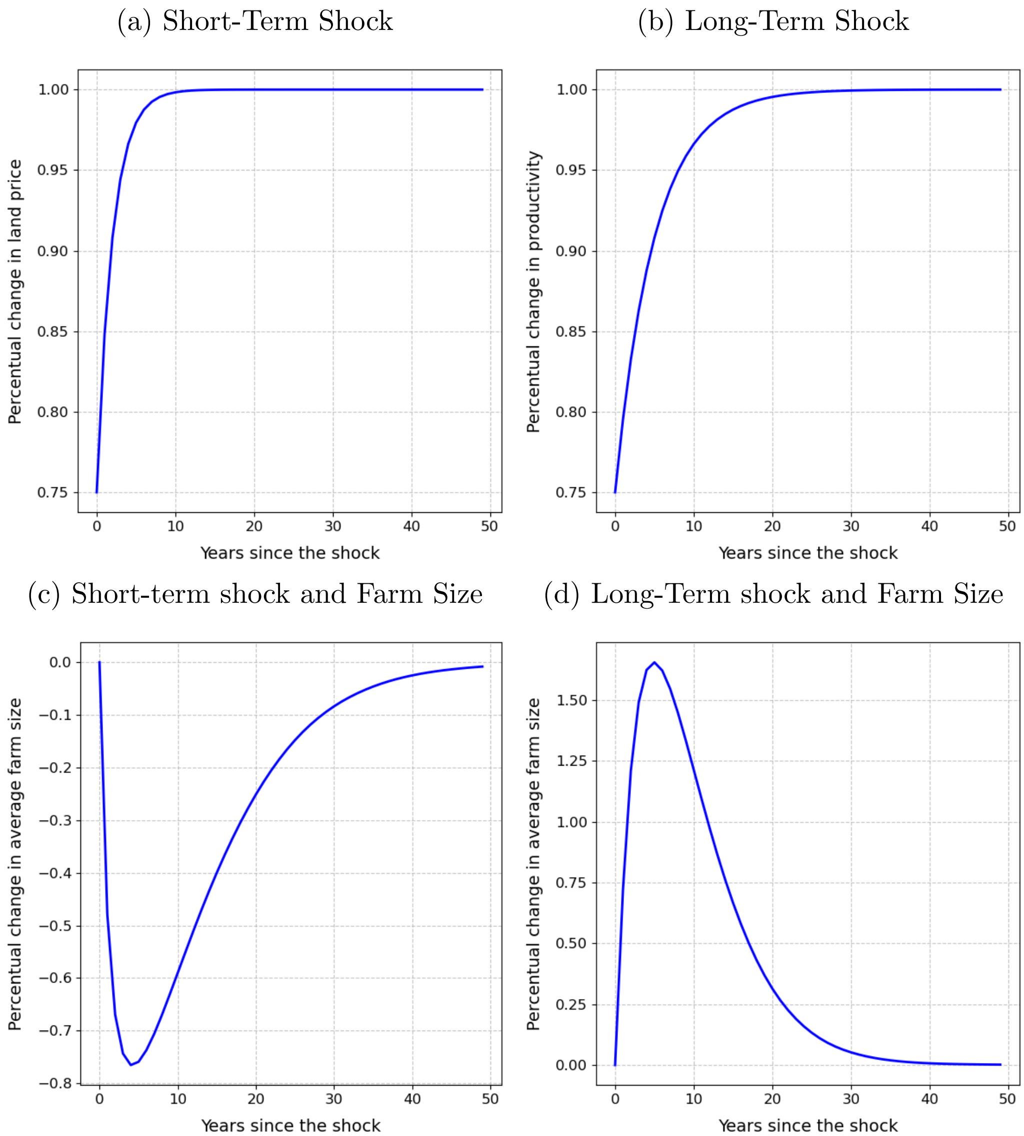
Notes: Data from the System of Information for Rural Development (SIDER)

Figure O.3: Ratio of Land Sales to Number of Allocations



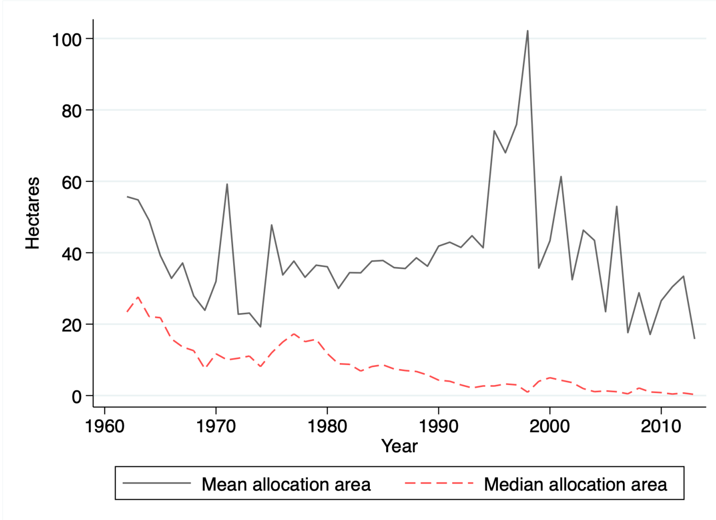
Notes: Data from the National Superintendence of Notaries (SNR). The figure shows the proportion of farms sold in each *vereda* to the total number of farms allocated by the government between 1980 and 2011.

Figure O.4: Agricultural Productivity Shock with Strong Persistence



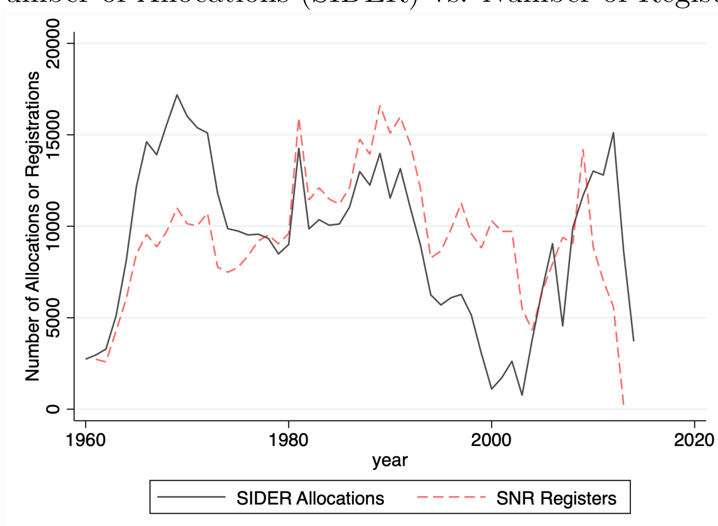
Notes: Panel (a) shows the simulated short-term shock and Panel (b) shows the long-term shock. Panel (c) and Panel (d) show the implications for the average farm size. In the short-term case, we have an increase in the number of farmers, reducing average farm size, whereas in the long-term case we observe an increase instead, with the exit of farmers being larger than the inflow. In both cases, the long-term implications are substantial.

Figure O.5: Mean and Median Allocation Size - 1961–2012



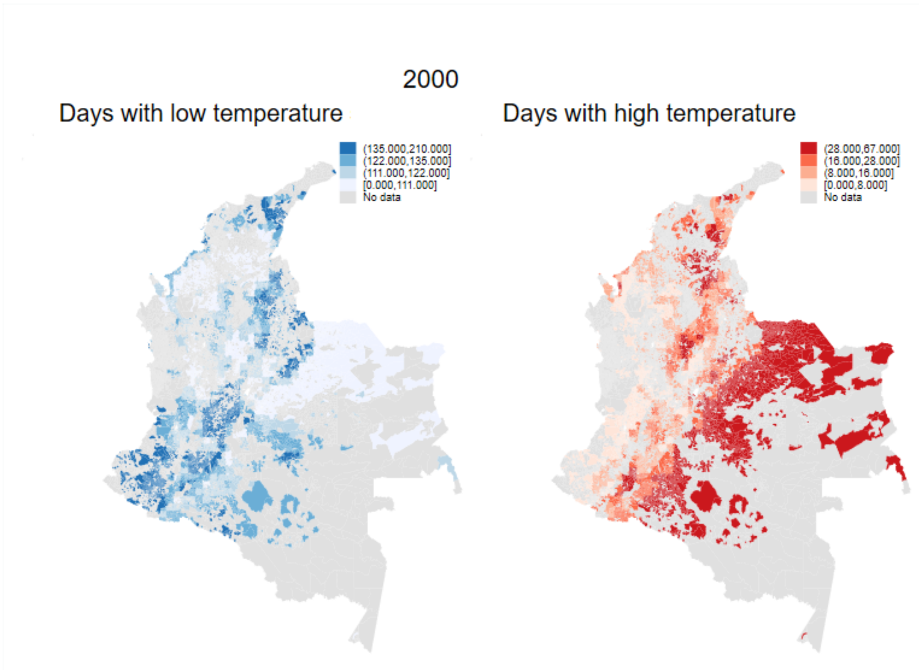
Notes: Data from the System of Information for Rural Development (SIDER). National-level yearly average area of land properties granted by the government as part of the public-land allocation program.

Figure O.6: Number of Allocations (SIDER) vs. Number of Registrations (SNR)

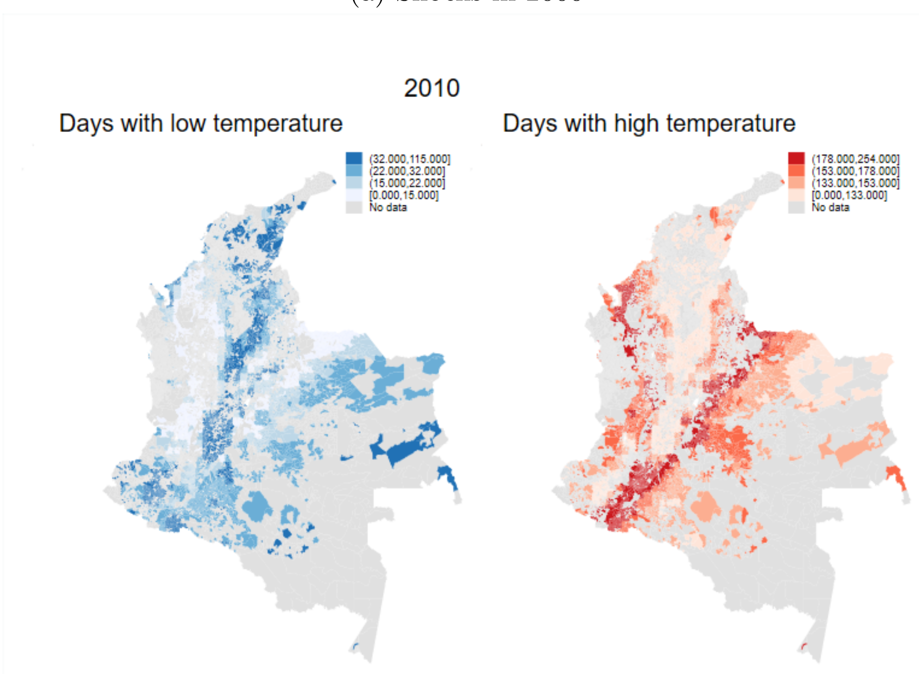


Notes: Data from the System of Information for Rural Development (SIDER) and from the National Superintendency of Notaries (SNR). The figure compares the number of land properties allocated by the government as part of the public-land allocation program with the number of properties registered at local 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 O.7: Temperature Shocks Across Space - 2000 and 2010



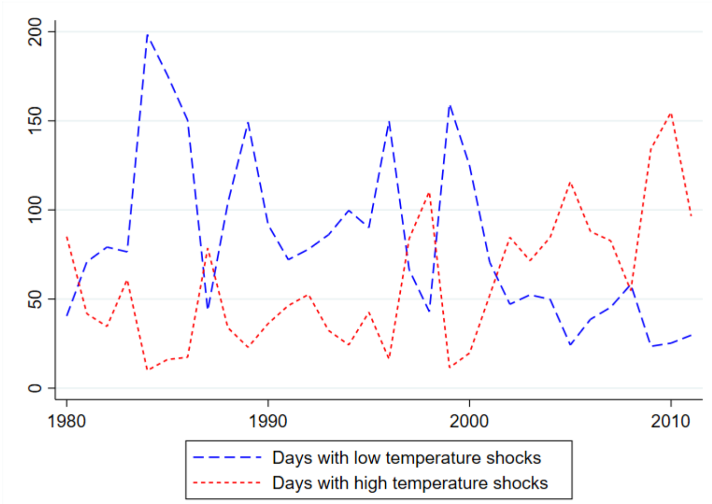
(a) Shocks in 2000



(b) Shocks in 2010

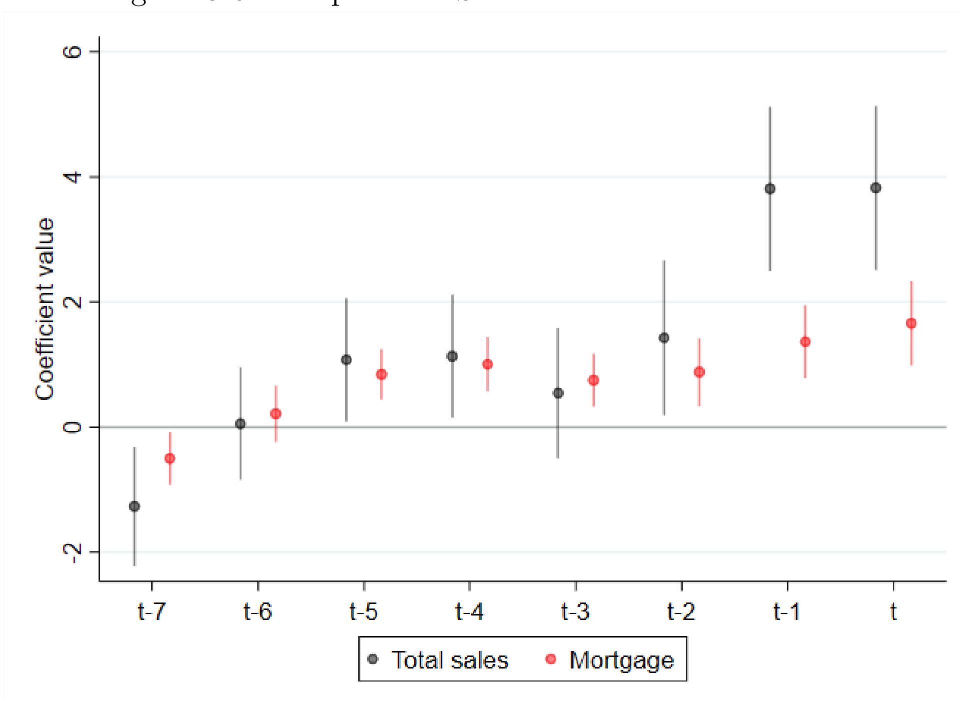
Notes: Data from the Copernicus Climate Change Service (C3S). The figure shows the average number of days with extreme heat (red) and cold (blue) across veredas in our sample in 2000 and 2010.

Figure O.8: 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.

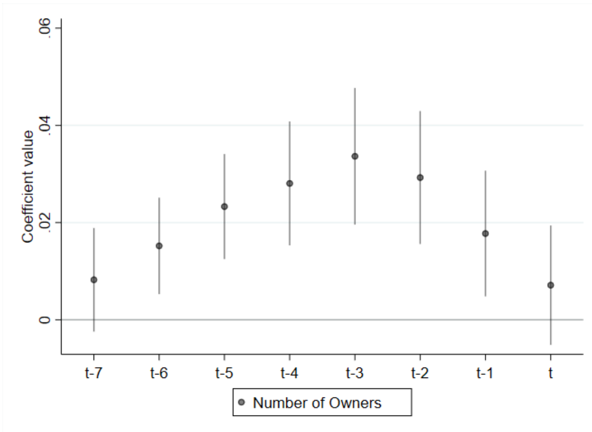
Figure O.9: Temperature Shocks and Land Transactions



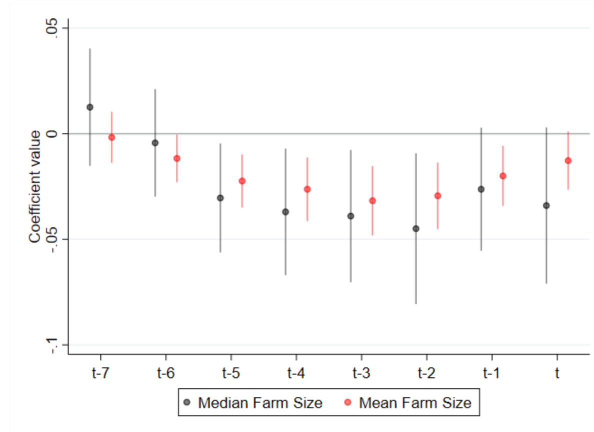
Notes: Data from the National Superintendency of Notaries (SNR) records. The figure presents the coeffiecient estimates for the number of days with atypical temperature (divided by 100) and its lags from equation (O.1). Dependent variables are: the total number of land sales (black) and the total number of land mortages (red). Included controls are accumulated allocations, accumulated precipitation at time t and five lags. Regressions also include year and municipality fixed effects Standard errors clustered at the municipality level. The vertical lines represent the 95% confidence intervals

Figure O.10: Temperature Shocks and Farm Size

(a) Owners

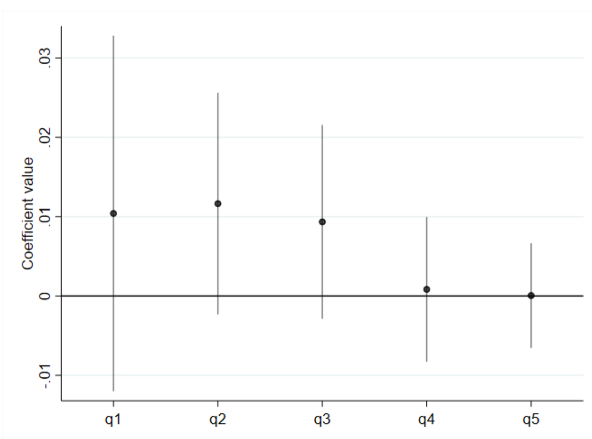


(b) Area

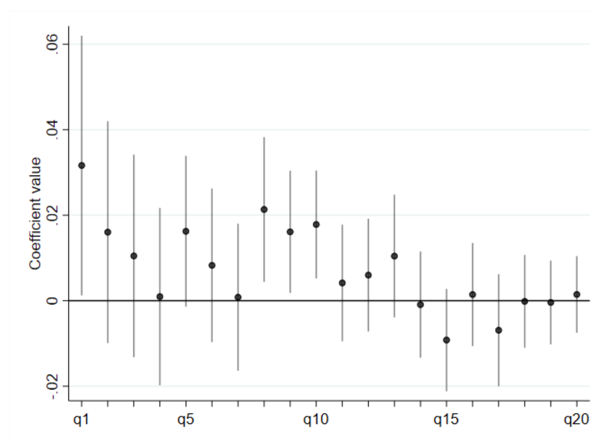


Notes: National Land Registry (*Catastro Nacional*), mantained by the National Geographical Institute (IGAC). The figure presents the coeffiecient estimates of the number of days with atypical temperature at between t and $t-7$ divided by 100 from equation (O.1) using as dependent variables the total number of land owners (panel a) and alternative measures of farm size (panel b). Included controls are log accumulated allocations, log area covered by the land registry, an indicator of registry updates, log accumulated precipitation at time t and its seven lags. Regressions also include year and municipality fixed effects. Standard errors clustered at the municipality level. The vertical lines represent the 95% confidence intervals.

Figure O.11: Temperature Shocks and Number of Owners by Initial Distribution Quantiles - Alternative Partitions



(a) Quintiles



(b) Vigintiles

Notes: OLS estimates of the γ coefficients according to equation (6), for each of the 5 and 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 ($t-1$, $t-2$) divided by 100. The dependent variable is the log number of owners per quantile. Controls are log accumulated allocations, log area covered by the land registry, an indicator of registry updates, log accumulated precipitation and five lags. Regressions also include year and municipality fixed effects. Error bars display 95% confidence intervals for standard errors clustered at the municipality level.

References

Acampora, M., Casaburi, L., and Willis, J. (2022). Land rental markets: Experimental evidence from kenya. Technical report, National Bureau of Economic Research.

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

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.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–684.

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, A. (2017). Fondo de Tierras del Acuerdo Agrario de La Habana: Estimaciones y propuestas alternativas. *Documento CEDE*, No. 41.

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2021). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. *Econometrica*, 89(5):2375–2408.

Avila, A. F. D. and Evenson, R. E. (2010). Total factor productivity growth in agriculture: The role of technological capital. *Handbook of agricultural economics*, 4:3769–3822.

Bolhuis, M. A., Rachapalli, S. R., and Restuccia, D. (2021). Misallocation in indian agriculture. Technical report, National Bureau of Economic Research.

Brooks, W. and Donovan, K. (2020). Eliminating uncertainty in market access: The impact of new bridges in rural nicaragua. *Econometrica*, 88(5):1965–1997.

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. 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.

Clementi, G. L. and Palazzo, B. (2016). Entry, exit, firm dynamics, and aggregate fluctuations. *American Economic Journal: Macroeconomics*, 8(3):1–41.

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

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

DANE (2019). Encuesta nacional agropecuaria, reporte tecnico. Technical report, Departamento Nacional de Estadística, Colombia.

de Roux, N. (2021). Exogenous shocks, credit reports and access to credit: Evidence from colombian coffee producers. *Documento CEDE*, (57).

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.

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.

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

Greenwood, J., Kircher, P., Santos, C., and Tertilt, M. (2019). An equilibrium model of the african hiv/aids epidemic. *Econometrica*, 87(4):1081–1113.

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*.

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and structural transformation. *Handbook of economic growth*, 2:855–941.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150.

Ibánez, A., Munoz-Mora, J., and Gafaro, 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.

Jaimovich, N., Terry, S. J., and Vincent, N. (2023). The empirical distribution of firm dynamics and its macro implications. Technical report, National Bureau of Economic Research.

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.

Kaboski, J. P., Lipscomb, M., Midrigan, V., and Pelnik, C. (2022). How important are investment indivisibilities for development? experimental evidence from uganda. Technical report, National Bureau of Economic Research.

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.

Khan, A. and Thomas, J. K. (2008). Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2):395–436.

Krishna, A. (2010). *One illness away: Why people become poor and how they escape poverty*. OUP Oxford.

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*.

Manysheva, K. (2022). Land property rights, financial frictions, and resource allocation in developing countries. Technical report, Working Paper.

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

Mazur, K. and Tetenyi, L. (2024). The macroeconomic impact of agricultural input subsidies. *Working Paper*.

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.

Pellegrina, H. S. (2022). Trade, productivity, and the spatial organization of agriculture: Evidence from brazil. *Journal of Development Economics*, 156:102816.

Peralta-Alva, A., Tam, X. S., Tang, X., and Tavares, M. M. (2023). Tax revenues in low-income countries. *The Economic Journal*, 133(653):2001–2024.

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.