Trade, Land Consolidation, and Agricultural Productivity*

Sebastian Sardon†

December 3, 2025

Job Market Paper

Updated Regularly. Link to the Latest Version.

Abstract

The agricultural sector features a large productivity gap between rich and developing countries, as well as substantial barriers inhibiting trade between them. In this paper, I assess whether removing such barriers can boost productivity in the developing world, and if so, through which channels. I study a 1997 USDA ban lift on avocado exports from the Mexican state of Michoacan to the U.S. Using a triple difference strategy and newly linked confidential census microdata (1970–2022), I find that the value of output per hectare increases by around 50% in treated areas suitable for avocado cultivation. Consistent with a model of agricultural production with heterogeneous farms, farms larger than 100 hectares roughly double their share of total land in treated areas. In the model, land consolidates because only large farms can afford the fixed costs of switching into avocados, which, unlike maize and other seasonal staples grown at baseline, require large amounts of water year-round. I use remote sensing to build a georeferenced dataset of irrigation investments and find a large investment response in suitable treated areas, on the order of 58% of Michoacan's baseline agricultural GDP. Finally, I find that these gains only occur in treated areas where land markets are not too frictional. In areas dominated by collective land tenure (*ejidos*), the effects of trade on investment, consolidation, and productivity all fail to materialize.

*I am deeply grateful to Lori Beaman, Christopher Udry, and Jacopo Ponticelli for the invaluable guidance and support I received throughout this project. I also thank Larry Christiano, Igal Hendel, Kiminori Matsuyama, Marti Mestieri, Prakash Mishra, Jose Ramon Morales-Arilla, Ameet Morjaria, Nancy Qian, Garima Sharma, James Sayre, Sebastian Sotelo, and Amilcar Velez for helpful feedback and suggestions. I benefited from comments received across multiple presentations in Northwestern's Development Breakfast and Lunch, the 2025 RIEF Annual Conference, and the MAP Seminar. This project would not have been possible without multiple travel grants from the Global Poverty Research Lab (GPRL) and a Small & PhD Research Grant from the Structural Transformation and Economic Growth (STEG) initiative. I thank Sean Higgins and Natalia Volkow for helping me access Mexican Census microdata, and I thank INEGI for extensive support to my registered project, *LM-2504*, which made this research possible. Finally, I benefited from field visits to avocado farms and conversations with stakeholders in the industry, including Alfonso Bustamante, Daniel Bustamante, Giovanni Cavaletto, Manuel Cerda, and Angeles Rubi. All errors are my own.

†Department of Economics, Northwestern University, Email: sebastiansardon2026@u.northwestern.edu

1 Introduction

Cross-country productivity differences are especially large in agriculture. Among countries in the bottom decile of GDP per capita, value added per worker in this sector is roughly 80 times lower than in the top decile (Restuccia et al., 2008; Gollin et al., 2014). A growing literature argues that misallocation of land across farms is an important source of this gap (Chen et al., 2023). In the poorest quintile of countries, farms smaller than 5 hectares cultivate 58% of planted land, compared to only 11% in the richest quintile, where average farm size is 34 times larger (Adamopoulos and Restuccia, 2014). Small farms also tend to be less productive than large ones (Foster and Rosenzweig, 2022). Policy efforts to raise agricultural productivity have mostly focused on smallholders through input subsidies, credit, and training programs. These "supply-side" interventions have failed to deliver substantial productivity gains in the aggregate (Suri and Udry, 2022). Recent work has argued for shifting attention to the demand side of markets (Goldberg and Reed, 2023), for example by lowering regulatory barriers to international trade. But whether such market-access reforms can meaningfully raise productivity in developing-country agriculture, and through which channels, remains largely unknown.

In this paper, I study a 1997 policy experiment in which, as part of the 1990-1994 NAFTA negotiations, the United States Department of Agriculture (USDA) partially lifted a century-old ban on Mexican avocado exports, triggering an export boom. Crucially, the initial ban lift was limited to one of Mexico's 32 states, Michoacan, generating sharp and exogenous variation in exposure to this trade shock.[1] Moreover, avocado cultivation is only viable under specific agro-climatic conditions that are not found everywhere within Michoacan. Access to a foreign avocado market is only valuable in areas suitable for avocado cultivation, generating within-state variation in exposure to the ban lift. Exploiting both sources of variation, I implement a triple difference design that compares outcomes in areas of Michoacan with high agro-climatic suitability for avocados (henceforth, "treated areas") against (i) unsuitable areas within Michoacan and (ii) suitable areas in neighboring states. This strategy controls for state-specific time-varying factors and time-varying confounders correlated with avocado suitability. I find that the ban lift significantly increased agricultural productivity in treated areas, with the value of output per

[1] The ban remained in place for all other 31 Mexican states until 2022, when it was lifted for the state of Jalisco. This policy of granting Mexico only partial access to the U.S. avocado market reflected a compromise after fierce negotiations with American avocado farmers (Stanford, 2002). The states of Nayarit and Mexico are obtaining ban lifts over the course of 2025.

hectare rising by over 50% relative to baseline.

To guide the analysis of the shock's mechanisms and heterogeneous effects, I develop a model of heterogeneous farms following Adamopoulos and Restuccia (2014). Building on an active literature on the reallocative effects of trade,[2] the model features trade opportunities that disproportionately benefit larger farms, which are the only ones able to afford the fixed costs required to participate in export markets (Melitz, 2003; Gáfaro and Pellegrina, 2022). This leads to land consolidation and productivity growth. However, such reallocations may be difficult in environments where land market transactions are costly or outright prohibited. To reflect these features of developing countries' land markets, I extend the model by introducing a collective farm sector. Collective farms control a fixed share of the land endowment, which is divided amongst members with an egalitarian rule and cannot be traded. Under fairly general conditions, the model predicts that trade should trigger land consolidation and lumpy private investments (the fixed costs), but that these effects should be weaker when the share of land controlled by collective farms is higher. Consistent with these predictions, I find evidence of land consolidation toward larger farms and a surge in irrigation investments. Moreover, these effects only materialize in areas where land markets function: in areas dominated by Mexico's collective land tenure system (*ejidos*), the trade shock has no significant effects.

The empirical analysis combines agricultural census data and government statistics with novel remote-sensing measurements of private irrigation investment, covering the 1970–2024 period. My main outcomes, measured at the municipality and enumeration area (EA)[3] levels, include the value of output per hectare at constant prices ("aggregate yield"), a measure of land concentration given by the share of land planted by farms with 100 or more hectares, and a remotely-sensed measure of lumpy irrigation investments in agriculture (irrigation ponds). As in De Haro (2022), I measure avocado suitability comparing statistical areas' climate and altitude with the FAO's ECOCROP tables on avocado's agronomic requirements.

This analysis reveals, first, that lifting the ban on Michoacan's avocado exports to the U.S. raised the value of output per hectare (aggregate yields) by over US$ 1,000 per hectare in treated areas,[4]

[2] See for example Pavcnik (2002), Trefler (2004), McCaig and Pavcnik (2018), and Felix (2021) for empirics, as well as Melitz (2003) and Melitz and Redding (2015) for theory.

[3] EAs are segments of municipalities defined by INEGI (*áreas geoestadísticas básicas*). They are on average 5 times smaller than a municipality.

[4] These results aggregate across crops using 1996 international price data from the FAO. Fixing prices at 1996 levels provides a conservative estimate of the gains from trade because, after the ban lift occurs (1997–2022), local avocado prices roughly double, net of overall inflation. The shock's effects on farm revenues are therefore even larger.

representing a 50% increase from baseline. These gains are driven by several mechanisms. They stem directly from a shift in land use: around 15% of planted land in treated areas shifts to avocados. Such land mostly comes from parcels originally devoted to maize, the crop occupying the majority of Michoacan's planted land. At prevailing prices and physical yields, avocados generate five to ten times more value per hectare than maize, so crop switching can mechanically generate the gains in output value documented above even without within-crop efficiency improvements. Indeed, I find that the gains are not driven by efficiency gains within specific crops: I can rule out positive effects on physical yields (metric tons per hectare) for avocado, maize, and other major crops. I present suggestive evidence that avocado yields remain stagnant because (i) the shock leads to farmers planting new orchards, which have lower yields than old ones, and (ii) avocado cultivation expands into less suitable land.

The switch of land from staple crops to avocado orchards is driven by land consolidation and this is in turn driven by farms needing lumpy private investments to make that switch. I find that the share of land planted by farms larger than 100 hectares doubles in response to the trade shock. This result is the first direct evidence of land reallocation in response to trade, and complements a prominent literature documenting firm reallocation in manufacturing (Bernard et al., 2007; Bustos, 2011; Atkin et al., 2017). In the model, trade shocks trigger land consolidation because switching entails large fixed costs, which are prohibitively expensive for small farms. Such fixed costs are difficult to measure directly, and a key contribution of this paper is to introduce a novel remote-sensing approach to measuring them in agriculture. I focus on irrigation investments required to set up avocado orchards, which, unlike the seasonal staples typically grown in the study area, require large amounts of water year-round. Private farms must bear these investment costs directly because of minimal public irrigation infrastructure in my paper's setting. Field interviews pointed to these investments, primarily in the form of irrigation ponds (*hoyas de agua*), as key barriers that prevent smallholders from being competitive in the avocado sector. Ponds are lumpy investments starting at around US$ 10,000—roughly equal to the mean annual revenue of farms in Michoacan.

To assemble a dataset that traces these investments over multiple decades, I leverage recent advances in remote sensing and computer vision that allow me to detect each irrigation pond in satellite imagery.[5] Following Couttenier et al. (2022) and Khachiyan et al. (2022), I train convolu-

[5] Irrigation ponds are visually distinct from surrounding cropland, and they are even easier to identify with infrared imagery using the normalized difference water index (NDWI).

tional neural network models on satellite imagery from Landsat-5 and Sentinel-2. The model used for the 2019–2024 period, where 10-meter high resolution Sentinel-2 imagery is available, achieves excellent performance (AUC = 0.992),[6] successfully identifying 95% of ponds while maintaining a false-positive rate below 30%. A related model used for 30-meter Landsat-5 imagery (1985–2015) achieves a more modest performance (AUC = 0.91, detecting 71% of ponds with a false positive rate around 45%). Together, the output of the two models allows me to build a dataset of irrigation investments over the 1985–2024 period. I find a statistically significant increase in private irrigation investment as a consequence of the U.S. avocado ban lift. This provides direct evidence on lumpy investments being a response to market access. In the model, such responses underlie the productivity and consolidation effects of trade. Back-of-the-envelope calculations based on publicly available pond construction estimates indicate total investment effects on the order of US$ 1.3 billion, or 58% of Michoacan's baseline agricultural GDP (1996).

In addition, using a novel panel of farms I created by matching census waves, I find direct evidence of efficiency-enhancing reallocation: initially more productive farms experience faster growth in avocado-suitable areas after the ban lift, while farm growth is independent of baseline productivity everywhere else. As predicted by the model, initially larger and more productive farms are more likely to start producing avocados, and farms making this crop switch expand their landholdings.

Three supplementary tests bolster the credibility of these results. As with other types of difference-in-differences designs, my results only have a causal interpretation under a parallel trends assumption. In particular, I need to assume that, in the absence of the trade shock, trends in outcome gaps between suitable and unsuitable areas would have evolved in a parallel manner in Michoacan and in neighboring states. In support of this assumption, I show that suitable-unsuitable outcome gaps evolved in parallel in the pre-treatment period (1970-1991). Second, I also rule out that these results are driven by large farms expanding into uncultivated land: the trade shock has no effect on treated areas' total planted land, so an expansion of large farms' land share must reflect land reallocation from smaller farms. Third, as a placebo test, I test whether non-avocado farms in treated areas experience consolidation. I find no increase in the share of land on non-avocado farms larger than 100 hectares. This assuages concerns about confounding

[6] Given a classifier, i.e. a continuous score indicating the likelihood of a given pixel containing a pond, an AUC of 0.992 means the classifier correctly ranks pixels with and without ponds 99.2% of the time, giving a higher score to the former. An AUC of 0.5 is achievable by random guessing, while an AUC of 1.0 is perfect classification.

factors driving differential land consolidation trends in treated areas.

In the last part of the analysis, I turn to the role of domestic frictions in shaping the gains from trade. A central theme in the literature on trade and development is that the gains from trade may fail to be realized when domestic frictions are too acute (Atkin and Khandelwal, 2020). In the agricultural sector, land market frictions may play a central role, as they make it difficult to buy and sell this critical input. In the limit, if land is essentially not tradable, then the reallocative effects of trade may fail to materialize, and the effects of trade on productivity may be muted. Mexico's dual land tenure system, with approximately half of agricultural land held in collective farms (*ejidos*), might generate frictions that could reduce the productivity impact of trade (De Janvry et al., 2015).

These findings have important policy implications. First, developed countries can directly contribute to agricultural productivity growth in the developing world by reducing their trade barriers. Though tariff barriers had been falling until recently, over 80% of agricultural products still face non-tariff barriers (Gourdon et al., 2015), and recent changes in the global policy environment suggest that both types of barriers may intensify over the next few years (Ignatenko et al., 2025). Second, developing countries stand to gain more from trade liberalization when it is accompanied by domestic land market reforms. Policies that facilitate land transactions and allow for farm consolidation can magnify the productivity response to new export opportunities. More broadly, the results point to important complementarities between reforms that liberalize output markets and those that liberalize input markets.

Related Literature This paper relates and contributes to several strands of literature. First, this paper relates to work on agricultural productivity and misallocation. A large literature documents that agricultural productivity severely lags behind the global frontier in poor countries and that differences in farm size and resource misallocation are central to these gaps (Restuccia et al., 2008; Eastwood et al., 2010; Gollin et al., 2014; Lagakos and Waugh, 2013; Adamopoulos and Restuccia, 2014, 2022; Chen et al., 2023). In contrast to developed economies, smallholders own the majority of agricultural land in developing countries and there is little evidence of cross-country convergence in farm sizes over time (Lowder et al., 2016, 2021). Despite substantial investment in subsidies, extension programs, and other policies targeted at smallholder farmers, aggregate agricultural productivity in the developing world has failed to converge to the frontier, with stagnation in some regions and declines in others (Suri and Udry, 2022; Wollburg et al., 2024). This paper presents evidence on a possible alternative approach: demand-side interventions granting developing countries access to international cash crop markets. The large productivity effects documented here, on the order of 50% of baseline, echo recent arguments by Goldberg and Reed (2023), who use cross-country evidence to highlight the importance of international market integration for economic growth.

The paper also introduces a compelling new natural experiment to the large literature studying the effects of trade on productivity. A recurrent theme in this literature, motivated by trade models with heterogeneous producers, is that trade access raises aggregate productivity via selection and reallocation, especially when fixed export costs screen out small, low-productivity firms (Melitz, 2003; Melitz and Redding, 2015). Empirically, the causal evidence on this mechanism comes from manufacturing (Pavcnik, 2002; Bernard et al., 2007; Bustos, 2011; Atkin et al., 2017; McCaig and Pavcnik, 2018), and this paper is the first to study it in the agricultural sector. In doing so, I also contribute to the smaller but growing literature on the effects of international trade on agriculture, which has been predominantly structural (Costinot and Donaldson, 2016; Sotelo, 2020; Gaigne and Gouel, 2022; Farrokhi and Pellegrina, 2023).

Finally, this paper contributes to the literature on land market institutions and development (Goldstein and Udry, 2008; Manysheva, 2022; Adamopoulos et al., 2022). It is particularly related to De Janvry et al. (2015), who show that Mexican ejidos hinder agricultural productivity and structural transformation. My results indicate that the distortions induced by such collective tenure systems may be particularly severe during trade liberalization episodes.

The rest of the paper is organized as follows. Section 2 describes the institutional setting. Section 3 presents the theoretical model. Section 4 describes the various datasets I built for the empirical analysis. Section 5 presents the empirical strategy. Section 6 presents the main results. Finally, Section 7 concludes.

2 Institutional Setting

The USDA's 1997 authorization of fresh Hass avocado exports from one state of Mexico to the U.S. has desirable features to study the effects of international market access on agricultural development. First, the program was limited to a single Mexican state (Michoacan), creating a relatively rare natural experiment.[7] Second, the shock hit a sector that already accounted for

[7]The typical USDA ban lift is applied to entire countries at once. See for example the avocado ban lifts for Peru (2010) and Colombia (2017) or the tangerine ban lifts for Chile (2004).

roughly one-third of Michoacan's agricultural GDP at baseline, and effectively doubled its potential market size. Third, it happened in a relatively data-rich environment in which agricultural censuses measure both inputs and outputs used by the universe of farms. The rest of this section describes the ban lift, places it in a global context, and discusses key features of Mexican agriculture that shape how farms adjust to trade.

2.1 The ban lift on avocado exports from Mexico to the U.S.

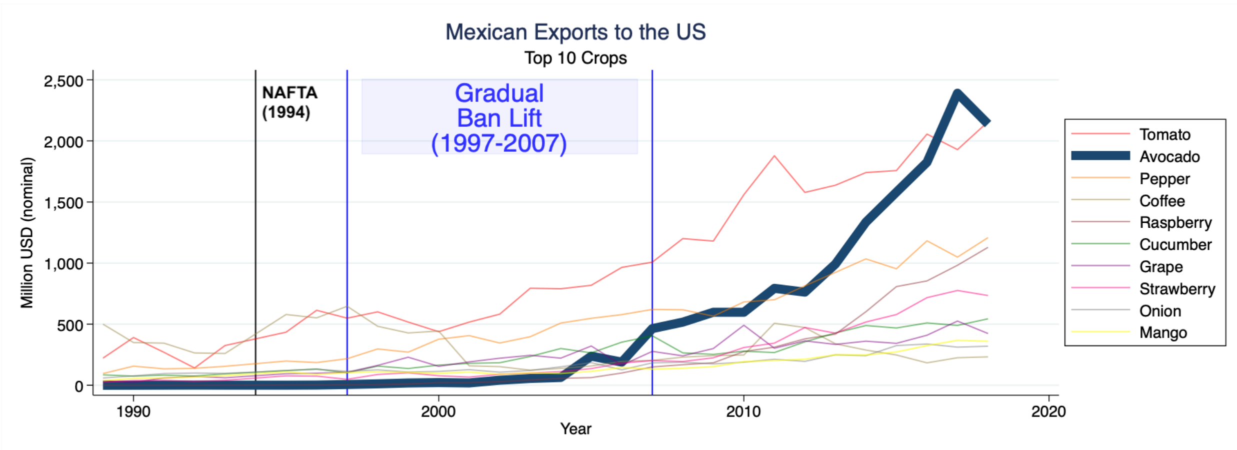
The United States Department of Agriculture (USDA) banned Mexican avocado exports from entering the United States in 1914 after American plant health officials found avocado seed weevils in Mexican orchards. Although officially phytosanitary, the decades-long ban has been characterized as covert protectionism driven by producer capture of the policy process (Grundke and Moser, 2019). And despite multiple attempts by Mexican authorities over subsequent decades, the ban persisted throughout the 20th century. NAFTA-related trade negotiations in 1990 led to the reconsideration of this longstanding policy (Bredahl, 2001).

On February 5, 1997, the USDA's Animal and Plant Health Inspection Service (APHIS) issued a final rule (effective March 7, 1997) permitting the importation of fresh Hass avocados grown in Michoacan, subject to strict phytosanitary safeguards.[8] The rule followed an intense period of negotiation involving USDA, Mexico's Secretariat of Agriculture (SAGARPA), and the state government of Michoacan, as well as representatives of both American and Mexican avocado farmers. The final agreement involved a compromise with American avocado producers, who opposed full liberalization of the U.S. market. The ban was only lifted for avocados produced in the state of Michoacan. Moreover, the ban was lifted gradually, over a 1997-2007 period. Michoacan started by obtaining, in 1997, access to Alaska and 19 Northeastern U.S. states during the months of winter. Market access gradually expanded to year-round and to all U.S. states over the ten years that followed.[9] Figure 1 illustrates the dramatic impact of this policy change: while Mexican avocado exports to the United States were negligible before 1997 (less than US$ 50 million annually), they grew steadily during the gradual liberalization period and accelerated sharply after full market access was achieved in 2007.

[8] 62 Fed. Reg. 5293 (Feb. 5, 1997), "Importation of Fresh Hass Avocado Fruit Grown in Michoacan, Mexico."

[9] In November 2001, APHIS lengthened the shipping season (to roughly October 15-April 15) and added destination States, while retaining the Michoacan-only origin restrictions. In November 2004, APHIS authorized year-round distribution nationwide but delayed entry into California, Florida, and Hawaii for two years. By early 2007, Mexican Hass avocados from approved municipalities in Michoacan could be distributed in all 50 States.

Figure 1: Mexican agricultural exports to the U.S.



Notes: This figure shows Mexican agricultural exports to the U.S. between 1988 and 2022, using data from UN Comtrade. The bold blue series shows total avocado exports, while the other series show exports for the other Mexican agricultural products. The figure focuses on the top 10 crops with highest export value in 2022. The black vertical line in 1994 indicates the year NAFTA came into effect, while the two blue vertical lines in 1997 and 2007 mark the beginning and end of the gradual ban lift on Mexican avocado exports to the U.S. By the end of 2007, Mexican avocados from Michoacan had full access to the U.S. market, conditional on complying with phytosanitary regulations.

Despite obtaining access to the U.S. market, Mexican exporters had to comply with strict regulations. Only some municipalities would be authorized for avocado exports, and they would need to be surveyed at least annually for pests. Orchards and packinghouses had to meet sanitation, inspection, and registration requirements (e.g., periodic removal of fallen fruit). Each shipment required appropriate labeling indicating the orchard of origin of every single fruit in it, as well as the identity of the grower, packer, and exporter. After packing, fruits had to be loaded into refrigerated containers that remained unopened until arriving in the U.S. Violations could trigger various penalties, ranging from the suspension of exports from the orchard of origin to the banning of an entire municipality.

Throughout my study window around the 1997 ban lift, origin eligibility was confined to Michoacan. The first additional state to obtain eligibility was Jalisco, in 2022. Nayarit and the State of Mexico are obtaining eligibility in 2025. My main study sample includes Michoacan and its six neighboring states, which include Jalisco and State of Mexico but not Nayarit. Because Jalisco first becomes eligible in 2022, any resulting adjustments to trade there would bias my long-run estimates toward zero, so the inclusion of Jalisco in the control group is conservative.

2.2 Global perspective: related ban lifts and non-tariff trade barriers

The lifting of the U.S. ban on Mexican avocados is but one example of multiple non-tariff barriers to agricultural trade that have been lifted over the last decades. In 2010, USDA granted a similar ban lift to Peruvian Hass avocados under similar conditions to the ones imposed on Michoacan.[10] Colombia followed in 2017 with essentially the same requirements.[11] By 2022, the United States imported US$ 264 million from Peru and US$ 44 million from Colombia in avocados.[12] USDA has applied similar phytosanitary safeguards to other fruits. In May 2009, a final rule allowed imports of sweet oranges and grapefruit from Chile, subject to cold treatment and fumigation requirements.[13] That opening fueled a surge in Chilean citrus exports and led to negotiations that would grant similar access to table grapes from that country.

Outside the United States, many trade barriers persist—especially in the European Union, the largest market for African and South Asian exporters. Although more than 90 percent of African agricultural exports entered the EU duty-free in 2019, non-tariff measures remain highly restrictive.[14] These include stringent pesticide residue limits and, more recently, sustainability mandates. For example, the EU's 2025 deforestation-free supply rule requires verifiable proof that commodities such as cocoa, coffee and timber are "produced without deforestation". This paper sheds light on the untapped potential of similar reforms elsewhere: if rich countries eased barriers as the United States did for avocados, other producers could see investment booms and productivity gains analogous to those documented in Michoacan (Beghin et al., 2015; Rao et al., 2012).

2.3 Mexican agriculture: Frictional markets and lagging productivity

The theoretical framework that follows seeks to accommodate salient features of Mexican agriculture that shape how farms adjust to trade. A first empirical regularity is fragmentation: the sector is dominated by many smallholders and large farms are rare. At baseline (1991), most of the planted land (52%) in Michoacan and its neighboring states was owned by smallholder farms under 5 hectares. Meanwhile, farms larger than 100 hectares only held 7% of the land. Countries

[10] Federal Register (2010), "Importation of Hass Avocados from Peru," USDA APHIS final rule.
[11] Federal Register (2017), "Importation of Hass Avocados from Colombia," USDA APHIS final rule.
[12] UN Comtrade (via WITS), United States: imports of products with code HS 080440 by partner (Peru, Colombia), 2023.
[13] Federal Register (2009), "Importation of Sweet Oranges and Grapefruit from Chile," USDA APHIS final rule.
[14] Global Panel on Agriculture and Food Systems for Nutrition (2025), "Implications of EU regulations."

in the top quintile of GDP per capita exhibit a very different allocation of land across farms, with smallholders below 5 hectares only holding around 10% of land (Adamopoulos and Restuccia, 2014).

A second central feature is Mexico's dual tenure system. A large share of agricultural land belongs to *ejidos* (collective farms), where property rights have historically been incomplete and transactions restricted. In 1992, the government launched *PROCEDE*, which issued parcel certificates and created a legal path to convert certificates into private property titles tradable with members outside the *ejido*. Rollout was nationwide over 1992–2006 and covered the vast majority of *ejidos* (De Janvry et al., 2014). However, these reforms stopped short of fully liberalizing the land market. External sales of *ejido* parcels still required assembly approval, and the program was explicitly designed to limit the use of ejido land as collateral. Consistent with that design, early work found no improvement in credit access (De Janvry et al., 2015). Because of this, where collective tenure still dominates, land transactions remain severely constrained. This keeps a large fraction of land effectively non-tradable, raising the cost of reallocating land toward farms that could profitably adopt cash crops by paying lumpy fixed costs such as irrigation.

These frictions are not unique to Mexico. Communal or customary tenure is widespread across the developing world. In Sub-Saharan Africa, large shares of rural land are governed by customary rules: in Tanzania, about 70% of land is designated as customary and roughly 80% of the rural population lives on it, with formal plot documentation still rare (Manysheva, 2022). Cross-country variation is stark, with the share of planted land owned communally ranging from roughly 2% in Rwanda to 97% in Somalia. Outside Africa, collective or state-dominated regimes are also common. For example, public-land systems in China and Vietnam also make land difficult to buy and sell (De Janvry et al., 2014; Manysheva, 2022). Collective tenure arrangements are thus a pervasive feature of land markets in the developing world.

3 Theoretical Framework

I study Mexican agricultural productivity in partial equilibrium, abstracting from non-agricultural sectors and assuming that the only variable input used in production is land. Once the economy opens to trade, there will be two agricultural sub-sectors: a cash crop demanded by a foreign market (denoted by $s = A$, including e.g., avocados), and a staple (denoted by $s = B$, e.g., beans and maize). The economy starts in autarky, fully specialized in sector B. Then, it opens to trade, giving farms the chance to switch to sector $s = A$ by paying a fixed cost f_A. This two-sector model with heterogeneous producers is based on Gáfaro and Pellegrina (2022), and related to Hopenhayn (1992), Melitz (2003), Hopenhayn (2014), Segerstrom and Sugita (2015), and Bai et al. (2024).

After characterizing how trade affects the allocation of land across farms in a baseline efficient model without any frictions, I consider a richer formulation in which a share λ of the land endowment is subject to a friction which makes it non-tradable. This formulation captures in a stylized way how areas in which collective farms control a large share of the land endowment may face constraints in reallocating land efficiently in response to trade shocks.

3.1 Frictionless Model (benchmark)

There is a continuum of farmers with heterogeneous productivity φ drawn from some distribution $G(\varphi)$.[15] They each operate DRS technologies $y = \varphi\, \ell^{\alpha}$, where ℓ denotes planted land and $\alpha \in (0, 1)$ is a curvature parameter. I abstract from entry and exit—every farmer remains active throughout.[16] Land is in fixed supply equal to L.

3.1.1 Equilibrium in the Closed Economy

I start by solving the model under autarky, where sector A is inactive, so all farms operate in sector B and maximize profits $\pi_B(\varphi) = \varphi\ell_B(\varphi)^{\alpha} - w\ell_B(\varphi)$. Farm-level land demand in sector B then is:

$$\ell_B(\varphi, w) = \left(\frac{\alpha\varphi}{w}\right)^{\gamma}, \quad \text{where } \gamma \equiv \frac{1}{1-\alpha}. \tag{1}$$

A farm with productivity φ produces $y(\varphi) = \varphi^{\gamma}\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}$ units of the staple crop. And the land market clearing condition is $L = \int_{\varphi_{min}}^{\infty} \ell_B(\varphi, w)g(\varphi)d\varphi$. I define the productivity index $\widetilde{\varphi} \equiv \left(\int_{\varphi_{min}}^{\infty} \varphi^{\gamma} g(\varphi)d\varphi\right)^{1/\gamma}$ and define aggregate output $Y \equiv \int_{\varphi_{min}}^{\infty} y(\varphi)g(\varphi)d\varphi$. Then, the land market clearing condition can be re-written using (1) as

$$L = \left(\frac{\alpha \cdot \widetilde{\varphi}}{w}\right)^{\gamma} \tag{2}$$

[15] All I assume is that G has a support of $[\varphi_{min}, \infty)$, a well-defined density g, and satisfies the regularity condition $\int_{\varphi_{min}}^{\infty} \varphi^{1/(1-\alpha)} g(\varphi)d\varphi < \infty$, which guarantees that the economy's demand for land is finite. I also normalize the measure of farmers to 1.

[16] This is the main simplification I made relative to Hopenhayn (1992) and Melitz (2003). I only allow the economy to adjust along the intensive margin of landholdings.

Moreover, I show in Appendix A that $Y = \widetilde{\varphi}L^{\alpha}$, so the heterogeneous farm economy's aggregate land demand and output coincide with those of a representative producer with productivity $\widetilde{\varphi}$. The land market clearing condition can be inverted to obtain the equilibrium land price in autarky:

$$w_0 = \alpha \cdot \widetilde{\varphi} \cdot L^{-(1-\alpha)}. \tag{3}$$

3.1.2 Equilibrium in the Open Economy

In the open economy, farmers can either produce staples (B) or an exportable cash crop (A). Farms in sector A sell their output at the exogenous world price $p_A > 1$ and incur a fixed cost f_A, whereas those in sector B face a price normalized to 1 and no fixed costs. Farms choose the sector yielding higher profits, which are given by: $\pi_A(\varphi) = \max_{\ell} \; p_A \, \varphi \, \ell^{\alpha} - w\ell - f_A$ and $\pi_B(\varphi) = \max_{\ell} \; \varphi \, \ell^{\alpha} - w\ell$. An additional assumption embedded in this formulation is *one-dimensional heterogeneity*: the same φ_i draw governs farm i's productivity in both crops.

Farm-level land demand in sector A is:

$$\ell_A(\varphi, w) = \rho \left(\frac{\alpha\varphi}{w}\right)^{\gamma}, \quad \text{where } \rho \equiv p_A^{\gamma}. \tag{4}$$

Substituting (4) and (1) into the profit functions of sectors A and B, respectively, yields: $\pi_A(\varphi) = \rho\varphi^{\gamma}\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}(1-\alpha) - f_A$ and $\pi_B(\varphi) = \varphi^{\gamma}\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}(1-\alpha)$. Farms choose sector A if and only if $\pi_A(\varphi) \geq \pi_B(\varphi)$, or equivalently iff $\varphi \geq \varphi^*$, with

$$\varphi^* = \frac{f_A^{1-\alpha} w^{\alpha}}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}(\rho-1)^{1-\alpha}} \tag{5}$$

The market clearing condition in the open economy is $L = \int_{\varphi^*}^{\infty} \ell_A(\varphi, w) g(\varphi) d\varphi + \int_{\varphi_{min}}^{\varphi^*} \ell_B(\varphi, w) g(\varphi) d\varphi$. Let $R \equiv \frac{\int_{\varphi^*}^{\infty} \varphi^{\gamma} g(\varphi) d\varphi}{\int_{\varphi_{min}}^{\infty} \varphi^{\gamma} g(\varphi) d\varphi}$, which I refer to as the *autarky share of potential exporters*. R measures the share of land used, in autarky, by the set of farms whose productivity is high enough to switch to sector A once the economy opens to trade. In Appendix A, I use this definition and the land market clearing condition to solve for the price of land under trade:

$$w_T = \alpha\widetilde{\varphi}[1 + (\rho - 1)R]^{1-\alpha} L^{-(1-\alpha)}. \tag{6}$$

Proposition 1. *The price of land in the open economy, w_T, is greater than the autarky price w_0.*

Proof Comparing equations (3) and (6), the price of land under trade w_T exceeds the autarky price w_0 if and only if $[1 + (\rho - 1)R]^{1-\alpha} \geq 1$. This inequality holds because $\rho = p_A^\gamma > 1$ and $R \geq 0$. □

I can now show that trade reallocates land from smaller to larger farms, and from lower-productivity to higher-productivity farms.

Proposition 2 (Land consolidation)**.** *Suppose the economy opens to trade. Let φ^* denote the cut-off productivity level that separates farmers in sector A from those in sector B. Then:*

i. For every farmer with $\varphi < \varphi^*$ (the set that remains in sector B), optimal land use falls.

ii. For every farmer with $\varphi \geq \varphi^*$ (the set that enters sector A), optimal land use rises.

Hence the opening of the export market reallocates land from lower-productivity and smaller farms to higher-productivity and larger farms.

Proof Part (i) follows from Proposition 1: since $w^T \geq w^0$ and $\partial \ell_B / \partial w < 0$, land for any $\varphi < \varphi^*$ falls. For (ii),

$$\frac{\ell_A(\varphi, w_T)}{\ell_B(\varphi, w_0)} = \frac{\rho(\alpha\varphi/w_T)^\gamma}{(\alpha\varphi/w_0)^\gamma} = \rho\left(\frac{w_0}{w_T}\right)^\gamma = \frac{\rho}{1 + (\rho - 1)R} \geq 1,$$

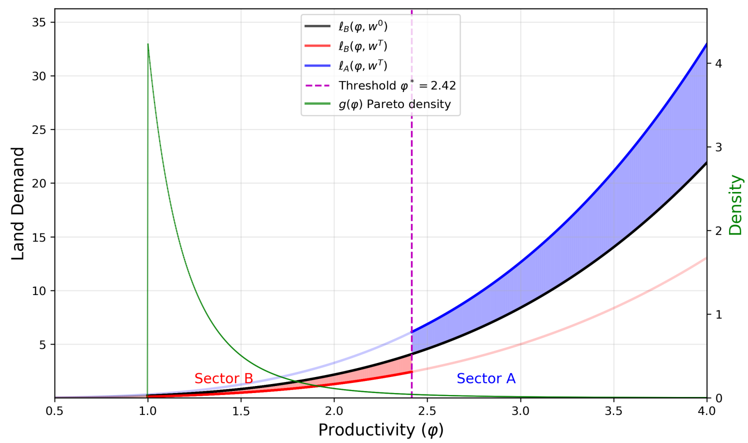
where the last equality uses (6). Thus, all farms switching to sector A expand. □

Figure 2 illustrates how Proposition 2 works in practice. The economy starts in autarky, fully specialized in sector B. Each farm's demand for land is given by the black curve, which corresponds to equation (1). The economy then opens to trade, and the price of land rises from w_0 to w_T (Proposition 1). In the open economy, the cut-off productivity φ^* determines whether farms switch to sector A or not. The red curve shows the new demand for land by farms who remain in B. These farms' demand still corresponds to (1) but is now lower due to land being more expensive. The blue curve shows the demand for land by farms who switch to sector A, who increase their landholdings.

3.2 Model with Land Market Frictions

Building on the frictionless benchmark, where trade induces land consolidation by reallocating resources to high-productivity farms, I now introduce land market frictions that closely mimic the land market of Mexico and other developing countries with collective landholding schemes.

Figure 2: Trade and Land Consolidation



Notes: The figure relies on the calibration in Table A1. The black curve shows land demand in autarky, where all farmers produce the staple of sector B. The red curve shows land demand in sector B after the economy opens to trade, while the blue curve shows land demand in sector A under trade. The vertical dashed line shows the cut-off productivity φ^* that determines whether farms in the economy choose sector A or B. The green curve shows the density of farm productivity in the economy, assumed to be Pareto for the illustration.

Specifically, I divide the total land endowment L into two regimes: a share $\lambda \in (0, 1)$ is non-tradable collective land (*ejido*), and the remaining share $1 - \lambda$ is freely tradable private land.

In the collective or "ejido" regime, land is allocated in fixed plots of size $\ell_c > 0$ to a measure $N_c = \lambda L/\ell_c$ of farmers, who cannot buy or sell land. An ejido farmer's profit in Sector B is $\pi_{B,c}(\varphi) = \varphi \ell_c^\alpha$, and in Sector A is $\pi_{A,c}(\varphi) = p_A \varphi \ell_c^\alpha - f_A$. When the economy opens to trade, those with $\varphi > \varphi_c^* = f_A/[(p_A - 1)\ell_c^\alpha]$ switch to sector A, but remain constrained to operating landholdings of size ℓ_c and cannot expand.

Farmers under the private regime have measure one and are identical to those studied in the benchmark model. They can trade land freely at a market price w, which clears within the private sector. Their crop choice and optimal land use follow the frictionless benchmark, with land demand curves $\ell_{A,p}$ for sector A given by (4) and $\ell_{B,p}$ for sector B given by (1). As in the frictionless model, sector choice is based on a cutoff φ_p^* that determines whether $\pi_{A,p}(\varphi_p^*) \geq \pi_{B,p}(\varphi_p^*)$. The private land price w equates total private demand for land to $(1 - \lambda)L$. Proposition 1 still holds: the price of private land rises under trade, $w_T > w_0$.

I next consider whether land market frictions limit reallocation. A measure of the amount of land that gets reallocated from low-productivity to high-productivity farms in response to trade is given by:

$$C(\lambda) = \int_{\varphi_p^*}^{\infty} \Big(\ell_{A,p}(\varphi, w_T) - \ell_{B,p}(\varphi, w_0)\Big) g(\varphi) d\varphi \tag{7}$$

In Appendix A, I show that $C(\lambda)$ can be rewritten as:

$$C(\lambda) = (1-\lambda) L R(\lambda) \left[\frac{(\rho-1)(1-R(\lambda))}{1+(\rho-1)R(\lambda)}\right] \tag{8}$$

Proposition 3 (Frictions Dampen Consolidation)**.** *Suppose the costs of switching to the export market are sufficiently high so that the share of farms that switches to sector A in the private regime owns at most half of the land under autarky, i.e.* $R < 0.5$*. Then, land consolidation is decreasing in the collective share* λ*, i.e.* $C'(\lambda) < 0$*.*

Proof See Appendix A.

Proposition 4 (Frictions Dampen Productivity Gains)**.** *If:*

(i) $\mathbb{E}[\varphi^2] := \int_{\varphi_{min}}^{\infty} \varphi^2 g(\varphi)\, d\varphi < \infty$ *(finite second moment),*

(ii) $\alpha > \frac{1}{2}$*,*

(iii) collective plots are fragmented, and, in particular,

$$\ell_c < \left[\frac{\alpha(1-\alpha)\rho^{-\alpha}(\rho-1)\widetilde{\varphi}\, f_A\, m}{(p_A-1)^2\mathbb{E}[\varphi^2]}\right]^{\frac{1}{2\alpha-1}} L^{\frac{\alpha-1}{2\alpha-1}},$$

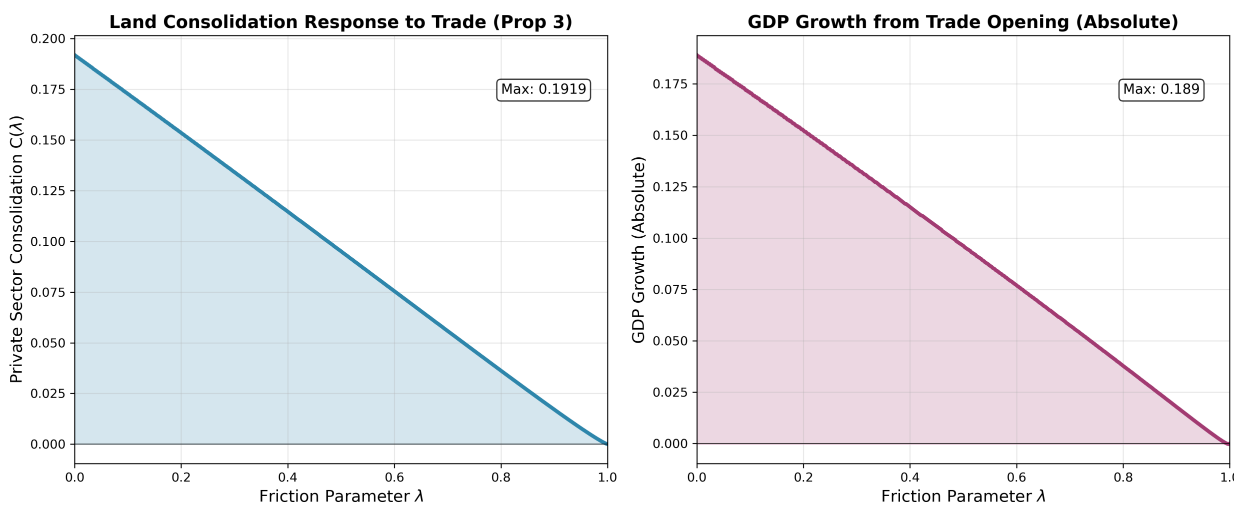
where $m \equiv \inf_{\lambda\in[0,1)}(1-\lambda)^{\alpha-1}R(\lambda)$*. Then, aggregate output gains from trade are strictly decreasing in the collective share* λ*.*

Proof See Appendix A.

The conditions imposed by Propositions 3 and 4 are rather loose. First, for Proposition 3 I need to rule out a situation in which most land is already in the hands of such large landowners, as this would already dampen the potential for a land consolidation response due to the trade shock and may lead to pathological cases in which, instead, the effect of frictions on the price of land dominates the effect of such frictions on dampening farms' ability to expand in response to trade. Second, for Proposition 4, I need to ensure that farms in the collective sector are not too big, and in particular the condition of this proposition helps rule out a case in which consolidated collective farms are *more* able than private farms to pay the fixed costs of switching to the export market and expand landholdings. With the parameter values chosen in Appendix A.6, the required threshold on ℓ_c is approximately equal to 0.21, meaning I need to rule out a situation in which a single collective farm owns over 20% of its statistical area's land. As argued in that Appendix, a plausible value for ℓ_c in this setting is on the order of 6×10^{-4}.

Figure 3 illustrates the comparative statics of Propositions 3 and 4, respectively. The figure shows how the consolidation measure $C(\lambda)$ and the output gain from trade $\Delta Y(\lambda) = Y_T(\lambda) - Y_0$ both decline monotonically as the collective land share λ increases from 0, representing the frictionless benchmark, to 1, representing a fully collective economy where no reallocation is possible. To facilitate interpretation, I plot $\Delta Y / Y_0$, expressing the output gains relative to initial output.

Figure 3: Comparative Statics: Frictions and Gains from Trade



Notes: The figure relies on the calibration in Table A1. The left panel shows how the consolidation measure $C(\lambda)$ in (7) declines as the collective land share λ increases. The right panel shows how the output gain from trade $\Delta Y(\lambda) = Y_T(\lambda) - Y_0$ also declines with λ. To facilitate interpretation, I plot $\Delta Y / Y_0$, expressing the output gains relative to initial output.

3.3 Model Discussion

The mechanism is straightforward. An export premium $p_A > 1$ (equivalently $\rho = p_A^{1/(1-\alpha)} > 1$) together with a fixed entry cost f_A generates a cutoff φ^* in (5). Opening to trade raises the private rental w_T relative to w_0 by the factor $[1 + (\rho - 1)R]^{1-\alpha}$ in (6), which tightens land for all farmers but raises the relative payoff to high–φ producers who clear the entry threshold. In the frictionless benchmark this produces consolidation: farms with $\varphi \geq \varphi^*$ expand and those with $\varphi < \varphi^*$ contract (Proposition 2), while a positive mass switches acreage from staples to the cash crop.

When introducing land market frictions, because collective plots cannot expand beyond ℓ_c, a higher λ limits the reallocative effects of trade. Under the restriction $R \leq \frac{1}{2}$, $C'(\lambda) < 0$ (Proposition 3) Frictions destroy the very margin—expansion by high–φ farms entering the avocado sector—that delivers gains from trade in this environment.

The model delivers some testable predictions that organize the empirical analysis that follows. High–productivity farms, which in the model are the only ones who can make this switch, should expand the share of the land endowment they own, and smaller units should reduce it. These patterns should attenuate where λ is high which in the empirical application is naturally captured by the share of land belonging to Mexico's *ejidos*. Finally, underlying this crop switch lie large fixed costs paid by large farms, which the remote sensing procedure described in the next section is designed to capture. These investments should also be attenuated where λ is high.

4 Data

Study Area and Overview of Sources The study focuses on the state of Michoacan, where the trade shock of interest occurs, and its six adjacent states (Colima, State of Mexico, Guanajuato, Guerrero, Jalisco, Queretaro). I use four main data sources: (i) official agricultural production statistics, (ii) Census data on the universe of farms, (iii) exogenous avocado suitability measures, and (iv) remote sensing of irrigation infrastructure. The data cover the 1970–2024 period.

Combining the four data sources mentioned, I build panel datasets at the farm, enumeration area (EA)[17] and municipality levels covering all of Michoacan and its six adjacent states (Colima, State of Mexico, Guanajuato, Guerrero, Jalisco, Queretaro). The Census permits building a municipality-level panel for 1970, 1991, 2007 and 2022, as well as farm- and EA-level panels for 1991, 2007 and 2022. Production statistics and remote-sensing measures are available, respectively, as municipality-year and EA-year panels over 2003–2022 and 1985–2022. The rest of the section describes each of the data sources in detail.

Administrative production data (SIAP). I construct a municipality-level panel on planted area and production by crop from Mexico's official agricultural statistics (SIAP, *Sistema de Información Agroalimentaria y Pesquera*). While state-level data are available from 1980, municipal coverage

[17] Enumeration areas, or *áreas geoestadísticas básicas*, are the lowest geographical units at which microdata can be extracted from INEGI's office. These areas are statistical but not administrative subdivisions of municipalities, with the typical municipality containing 5-10 enumeration areas.

begins in 2003. The data are collected through SIAP's monthly *Avance de Siembras y Cosechas* program, in which field technicians record hectares planted (*sembradas*) and tonnage harvested by crop and agricultural cycle. District offices review entries, and state delegations validate the data before releasing preliminary and definitive annual figures (SAGARPA, 2018).

My main analysis uses FAO world prices from 1996, the year before Mexico's avocado export liberalization, to construct crop-specific output values in 1996 US$. As robustness checks, I test alternative years and also use crop-specific farmgate prices (*precio medio rural*) from SIAP—the price paid to producers at first sale in the production zone. These farmgate prices are compiled from technicians' municipal estimates and aggregated to state-year series following SIAP's validation protocols.

The main outcome of interest in this paper is the aggregate value of agricultural production by each statistical area (EA or municipality). To build this outcome I use national base-year producer prices P_c^{base}, fixed in the pre-shock year of 1996, and compute

$$Y_{ist} = \Big(\sum_c P_c^{\text{base}} Q_{cist}\Big) \Big/ \Big(\sum_c L_{cist}\Big) \tag{9}$$

where Y_{ist} is the value of agricultural production per hectare in statistical area i (EA or municipality) in state s and year t, Q_{cist} is the quantity produced of crop c, and L_{cist} is the land planted with crop c. I treat missing quantities as zeros, though extensive verification with Census microdata reveals that the coverage of SIAP is complete. To enhance the precision of estimates, I winsorize outcomes at the 99% level. Results are robust to leaving the raw series untrimmed, but EA level values of Y_{ist} become extremely volatile when not winsorizing the data due to EAs where total planted land (the denominator of equation 9) is close to zero. Because municipality-year SIAP microdata are not available pre-2003, I recover baseline municipality production and land use levels for 1970 and 1991 from the Agricultural Census, which I describe next.

Agricultural Census Confidential farm-level microdata are available for 1991, 2007, and 2022 and are consulted on-site at INEGI. Each record is a production unit (farm) georeferenced to the *área geoestadística básica* (referred to as EA, enumeration area) and municipality and then aggregated at these levels for extraction from INEGI premises. I harmonize codes across waves using official crosswalks. When EAs split or merge, I reweight observations by farm counts and cultivated area to maintain a consistent EA definition.

At the farm level, I build a panel across 1991, 2007, and 2022. Linkage is based on fuzzy matches on farmer names, run separately for each municipality. I retain only high-confidence links[18], resulting in a match rate of 39% for 1991–2007 and 43% for 2007–2022. The resulting files report land operated (owned and rented; private vs. *ejido*), crop mix and quantities, sales and marketing, labor (family and hired), and technology (irrigation and other inputs). Earlier historical waves (1960-1981, print volumes) are being digitized for context but are not used in the main estimates.

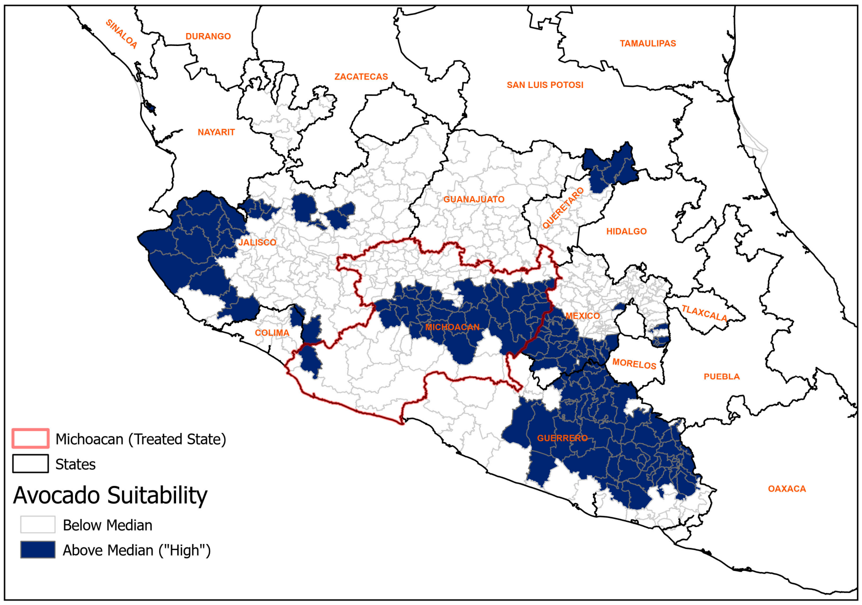
Avocado suitability. A central challenge in identifying the causal effects of the trade shock is defining exposure to the ban lift in a way that is exogenous to economic outcomes. Using actual avocado production to define treatment would be problematic: where avocados are grown reflects endogenous responses to prices, market access, and unobserved local economic conditions. Instead, I construct a measure of agronomic suitability for Hass avocado cultivation based exclusively on pre-determined agro-climatic conditions. The measure indicates whether a given statistical area *could* profitably grow avocados given its natural endowments. Only in areas suitable for avocado cultivation does the lifting of the U.S. import ban create valuable new export opportunities, generating exogenous variation in exposure to trade.

I follow De Haro (2022) by relying on the ECOCROP suitability model developed by agronomists (FAO, 2016). Figure 4 displays the resulting spatial pattern of avocado suitability across Michoacan and its six neighboring states. The ECOCROP model evaluates whether local climatic conditions meet the known agronomic requirements for a specific crop. For Hass avocados, these requirements are particularly stringent: the trees require moderate temperatures year-round (optimal range: 15-25°C), consistent precipitation or irrigation (minimum 800mm annually), and specific elevation ranges (typically 1,000-2,400 meters in central Mexico) where temperatures remain stable. Using daily temperature and precipitation data from NASA's AgMERRA dataset and altitude data from INEGI, I compute a continuous suitability index for each municipality that ranges from zero

[18]Concretely, I compute a bigram-based Jaccard score between the farmer names that exist in different waves. Bigrams are two-character subsets of a string variable. For example, the word *example* has the 6 bigrams *ex*, *xa*, *am*, *mp*, *pl* and *le*. Consider two farmer names found in different census waves, $name_1$ and $name_2$, each containing s_1 and s_2 bigrams respectively. Let m denote the number of bigrams contained in both names. Then, a Jaccard score for these two names is given by $m/\sqrt{s_1 s_2}$. This score is a measure of string similarity that ranges from 0 (completely different) to 1 (exact match). When linking the 1991 wave to the 2007 wave, and the 2007 wave to the 2022 wave, I compute this score for all possible pairs of names that include a name from each wave. I keep pairs with a score above 0.75 ("matches") and, when a farmer name gets matched to multiple names from the subsequent wave, I keep the match with the highest score. This procedure is implemented using Stata's `matchit` command and required approximately one week to run on INEGI's facilities.

(unsuitable) to one (optimal conditions). Critically, this index is constructed using only exogenous geographic and climatic variables.

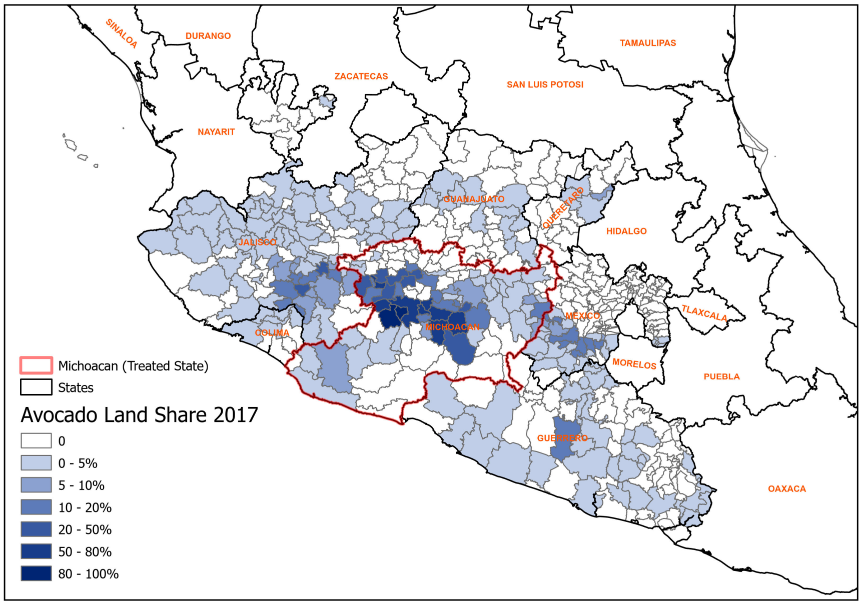
Figure 4: Avocado Suitability Across Michoacan and Neighboring States



Notes: Map displays agronomic suitability for Hass avocado cultivation based on altitude, temperature, and precipitation using the ECOCROP model (De Haro, 2022, Working Paper). Dark blue areas indicate above-median suitability ("High"). Michoacan (red outline) is the treated state that gained U.S. market access in 1997.

For my main analysis, I define high suitability as above-median values of this continuous index within my seven-state sample, creating a binary indicator $H_{is} = 1$ for treated areas. This binarization maintains substantial variation both within Michoacan (comparing suitable to unsuitable areas) and across states (comparing Michoacan's suitable areas to similar areas in neighboring states). The measure performs well empirically. Figure 5 shows the share of land planted with avocado in each municipality in year 2017, ten years after the 1997-2007 gradual ban lift of avocados from Michoacan had been fully implemented. With some exceptions in the state of Jalisco, the vast majority of municipalities specialized in avocado cultivation are classified as having high suitability by the H_{is} indicator. In first-stage regressions, presented in Table 1 for a cross-section of municipalities and EAs, suitability strongly predicts actual avocado cultivation, with an F-statistic > 50 when regressing municipality- and EA-level avocado cultivation on suitability. The table relies on statistical areas' *mean* share of land devoted to avocado cultivation over the 1970–2022 period, for municipalities, and over the 1991–2022 period, for EAs.

Figure 5: Actual Avocado Land Share in 2017



Notes: Map displays the share of agricultural land devoted to avocado cultivation in 2017 at the municipality level. Darker shades indicate higher avocado land shares, ranging from 0% (white) to 80-100% (darkest blue). The strong correspondence between this realized pattern and the predicted suitability in Figure 4 validates the use of agro-climatic suitability as an exogenous measure of exposure to the 1997 trade liberalization.

Remote sensing of irrigation investment. A central theme of my paper is that access to the U.S. avocado market disproportionately benefits the largest and most productive farms because they are the only ones capable of paying the high fixed costs associated with avocado production and commercialization. A key cost of this nature is the investment in irrigation infrastructure, which is essential for avocado cultivation because the crop requires a consistent water supply throughout the year while precipitation is highly seasonal in Michoacan. Since the government provides little to no support for irrigation infrastructure, farmers must finance these investments privately. I therefore expect to find that (i) access to the U.S. avocado market leads to an increase in private agricultural investment and (ii) in areas in which land market frictions (*ejidos*) are present, consolidation will not occur, less investment in ponds will be observed, and productivity growth will be lower. These irrigation ponds represent lumpy investments with substantial fixed costs. Using data from Mexico's government procurement database (Compranet), which records 32 geomembrane purchases for agricultural water storage between 1997 and 2024, the average cost for pond construction is approximately US$ 62,528, with costs ranging up to US$ 100,000 for larger ponds. For context, these average cost represents roughly twice the mean yearly output value

Table 1: First Stage: Suitability Predicts Avocado Cultivation

	(1) Avocado Share (%)	(2) Avocado Share (%)
Avocado Suitability	5.66***	4.14***
	(0.66)	(0.41)
F-stat	74.8	102.4
R2	0.164	0.29
N	489	2,416
Unit of Observation	Municipality	EA

Notes: The outcome variable is the share of planted land devoted to avocado cultivation. This outcome is regressed, exclusively, on a continuous agro-climatic suitability index for avocado cultivation, ranging from 0 (unsuitable) to 1 (highly suitable). The table presents "first-stage" regressions testing whether suitability indeed predicts specialization in avocado. Column (1) uses municipality-level data, while Column (2) uses EA-level data. In both cases, the outcome variable is the average of 2007 and 2022 cultivation shares, while the suitability index (conceptually, a static measure) is based on average weather over 1985-2015. Robust standard errors in parentheses. *** $p < 0.01$.

farms in Michoacan, making pond investment prohibitively expensive for smallholder farmers. To value the total investment represented by the detected ponds, I rely on cost estimates from local suppliers and official government procurement records, as described in Appendix Table B1.

To measure investments in irrigation ponds, I develop a two-stage machine learning algorithm that detects ponds in satellite imagery. The primary model is a convolutional neural network (CNN) that accurately identifies ponds in modern high resolution Sentinel-2 data (2019–2024). The secondary model is a smaller CNN that uses 2024 detections as reference points to determine if a pond existed at the same location in historical imagery, using lower-resolution Landsat data (1985–2018). Appendix B provides a detailed description of the algorithm.

The output of the algorithm is a series of pond presence predictions across the entire study period (1985–2024), which can then be collapsed at various levels for empirical analysis. For my main results, I aggregate data to the level of the enumeration area (EA)—the smallest geographic unit available in the Mexican census microdata.[19] I then use this dataset to estimate the causal effects of Michoacan gaining access to the U.S. avocado market on local investment patterns, and to test whether these effects are heterogeneous based on local land market frictions as proxied by the share of land belonging to collective farms (ejidos) in 1991, the Census wave prior to U.S. market access.

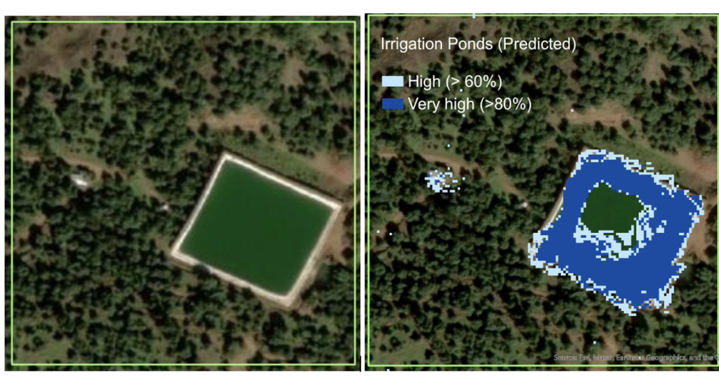
[19] Robustness tests will involve repeating the analysis at the coarser municipality level (the typical municipality has 5-10 EAs).

Currently, I have only built a dataset of remotely sensed ponds for the state of Michoacan. Because of this, I will present preliminary results using a difference-in-differences strategy that compares investment trends in avocado-suitable versus unsuitable areas within this state. In ongoing work, the neural networks trained in Michoacan will be deployed across its six neighboring states (Colima, Guerrero, Guanajuato, Jalisco, State of Mexico, and Queretaro). The expanded dataset will enable me to use the paper's main triple difference identification strategy.

Figure 6 illustrates the algorithm's detection capability, showing probability maps overlaid on high-resolution imagery that successfully identify pond boundaries. Table 2, together with Figures C1a and C1b in the Appendix, report validation performance for the pond detection models on held-out test data from Ario municipality, where manual labels enable rigorous evaluation.

The modern detection model achieves strong performance on contemporary Sentinel-2 imagery (AUC = 0.99), though the F1-score of 0.73 reflects that some inaccuracies do exist. Historical classification performance degrades with image resolution, as expected given the fundamental measurement challenges posed by coarser satellite data. The classifier shows substantial deterioration on 30-meter Landsat data from 1996 (AUC = 0.85, F1 = 0.64), where individual ponds often occupy only a handful of pixels. This performance degradation, while significant, does not preclude meaningful empirical analysis: even the lower-accuracy Landsat-based classifier produces sufficiently reliable pond presence indicators to measure long-run investment trends across the four decades preceding and following market liberalization.

Figure 6: Example of pond detection results



Notes: The figure shows an example of pond detection results applying the Sentinel-2 model on 2024 imagery. The left panel displays high-resolution Maxar imagery from Google Earth (0.5m resolution), while the right panel presents the corresponding pond probability map generated by the CNN. Darker areas indicate higher predicted probabilities of pond presence.

Table 2: Remote Sensing Model Performance

Model	Data Source	Validation AUC	Validation F1-Score
Modern Detection	Sentinel-2 (2024, 10m)	0.9920	0.7290
Historical Classifier	Landsat 5 (1996, 30m)	0.8540	0.6403

Notes: The table reports validation performance metrics for the pond detection models on held-out test data from Ario municipality, one of Michoacan's 113. The modern detection model uses high-resolution Sentinel-2 imagery from 2024, while the historical classifier uses lower-resolution Landsat 5 imagery from 1996. AUC refers to the Area Under the Receiver Operating Characteristic Curve, and the F1-Score is the harmonic mean of precision (probability that a pond prediction is true) and recall (probability that a pond in the validation set is captured in the model predictions).

Descriptive Statistics Table 3 presents descriptive statistics for the key variables used in the analysis, including mean values and standard deviations. These are computed separately for the group of treated areas and for the three control groups involved in the triple difference design that the next section describes. The treated group, consisting of avocado-suitable areas in Michoacan, exhibits higher average agricultural productivity and a greater prevalence of irrigation ponds post-liberalization compared to control groups. This preliminary comparison underscores the potential impact of U.S. market access on local agricultural outcomes. Analogous tables at the EA level are included in Table D1 in the Appendix.

5 Empirical Strategy

The empirical strategy exploits the fact that the U.S. granted fresh Hass avocado market access to a single Mexican state (Michoacan) beginning in 1997, with access expanding through 2007. Exposure to this policy is valuable only where agro-climatic conditions allow profitable avocado cultivation. I therefore implement a triple difference design that compares outcomes in *avocado-suitable* parts of Michoacan against (i) less-suitable parts within Michoacan and (ii) avocado-suitable parts in neighboring states. The first difference controls for state-specific time-varying confounders (e.g. macroeconomic shocks affecting statewide GDP); the second one controls for time-varying confounders that differentially affect avocado-suitable places throughout the region (e.g., avocado suitability correlates with tomato suitability and tomatoes also experience a boom throughout the study period).

Let i denote statistical areas (municipalities in the annual SIAP panel; enumeration areas, EAs, in census panels), s states, and t years. Let $T_s \in \{0, 1\}$ indicate the treated state (Michoacan),

Table 3: Descriptive Statistics by Treatment Status (Municipality Level)

Treatment Status	All	Treated	Control	Control	Control
Treated State	-	Yes	Yes	No	No
Avocado Suitability	-	Yes	No	Yes	No
Panel A: Production and Land Use					
Aggregate Yield	1,822.08	2,439.90	1,642.22	1,448.84	1,963.41
	(1,932.86)	(2,559.79)	(1,585.87)	(1,086.67)	(2,196.27)
Total Hectares	7,141.94	6,984.34	6,720.02	5,821.23	8,133.32
	(7,291.34)	(5,556.65)	(5,273.07)	(5,676.28)	(8,700.50)
Total Output (MM USD)	13.46	23.90	12.89	9.01	14.07
	(23.08)	(42.80)	(18.74)	(12.24)	(22.03)
Share Avocado (%)	3.09	18.02	2.15	1.53	0.87
	(12.02)	(28.89)	(7.75)	(5.24)	(4.75)
Share Maize (%)	76.01	67.04	65.10	84.66	75.03
	(27.74)	(34.26)	(28.07)	(21.60)	(27.80)
Panel B: Land Consolidation					
Share Land Above 100 Ha. (%)	6.87	3.11	5.73	5.89	8.65
	(8.99)	(4.47)	(7.26)	(7.54)	(10.50)
Share Land Below 5 Ha. (%)	67.67	73.23	74.48	69.14	63.79
	(22.21)	(17.29)	(16.93)	(21.88)	(23.84)
Panel C: Suitability					
Avocado Suitability	0.37	0.72	0.26	0.68	0.11
	(0.32)	(0.22)	(0.09)	(0.20)	(0.13)
Panel D: Avocado Ponds					
Pond Coverage (%)	0.31	0.23	0.38	—	—
	(0.46)	(0.26)	(0.59)	—	—
Municipalities	515	56	57	158	244
Observations	11,233	1,228	1,247	3,444	5,314

Notes: Panel A variables: Aggregate Yield - Value of total production, aggregated across all crops, relative to total planted land (1996 USD/Ha.); Total Hectares - Total planted land (Ha.); Total Output (MM USD) - Total value of production in millions of 1996 USD; Share Avocado - Share of avocado in total planted hectares (%); Share Maize - Share of maize in total planted hectares (%). Panel B variables: Share Land Above 100 Ha. - Share of privately owned land in parcels above 100 hectares (%); Share Land Below 5 Ha. - Share of privately owned land in parcels below 5 hectares (%). Panel C variables: Avocado Suitability - Continuous measure of avocado suitability. Panel D variables: Pond Coverage - Percentage of agricultural area covered by avocado ponds (%). Treated state refers to Michoacan. Avocado suitability is defined as above median suitability (weighted by 1991 hectares) within Michoacan. Standard deviations in parentheses. Unit of observation: Municipality. Data years: 1970, 1991, 2003-2022 (Panel A production variables from agricultural census); 1991, 2007, 2022 (Panel B land consolidation from agricultural census).

and let $H_{is} \in \{0, 1\}$ indicate high avocado suitability for statistical area i (above-median index) constructed ex ante from NASA climate and INEGI elevation, following De Haro (2022) and ECOCROP agronomy (Section 4). Because H_{is} depends only on geography and long-run weather, it is predetermined and exogenous to contemporaneous outcomes. The annual administrative panel runs 2003-2022; the census panels provide levels in 1970, 1991, 2007, 2022.

Using 1991 as the omitted year (period h_0 in the estimating equation), the specification is:

$$Y_{ist} = \alpha + \sum_{h \neq h_0} \beta_h^T \Big(T_s \times \mathbb{1}[t = h]\Big) + \sum_{h \neq h_0} \beta_h^H \Big(H_{is} \times \mathbb{1}[t = h]\Big) + \sum_{h \neq h_0} \beta_h^{TH} \Big(T_s \times H_{is} \times \mathbb{1}[t = h]\Big) + \theta_i + \theta_t + \varepsilon_{ist}, \tag{10}$$

where Y_{ist} denotes the outcome of interest, θ_i are statistical area fixed effects, and θ_t are year fixed effects. The coefficients $\{\beta_h^T\}$ absorb Michoacan-specific shocks in each year, $\{\beta_h^H\}$ absorb shocks that load differentially on avocado-suitable places in each year, and the triple interaction coefficients $\{\beta_h^{TH}\}$ capture the effects of access to the U.S. avocado market. Flat pre-1997 coefficients are a core identification check to rule out diverging pre-trends between treated and control areas.

The identifying assumption is that, absent the policy, the *suitability gap* (high minus low) would have evolved similarly in Michoacan and in the comparison states. Equation (10) enforces two layers of saturation that make this restriction plausible: (i) $T_s \times$year terms absorb the effects of any state-level program, enforcement, credit, or infrastructure shock common to Michoacan in year h, and (ii) $H_{is} \times$year terms absorb any regional shock that loads more heavily on avocado-suitable places in year h (e.g., climate cycles). The remaining identifying variation is the *difference* in the suitability gap between Michoacan and neighbors, over time. I verify the restriction with pre-1997 coefficients in the census panels and with re-centered pre-period checks in the annual series.

As a complement to the dynamic specification in (10), I will also estimate a simpler pooled triple difference specification:

$$Y_{ist} = \alpha + \beta^T \Big(T_s \times Post_t\Big) + \beta^H \Big(H_{is} \times Post_t\Big) + \beta^{TH} \Big(T_s \times H_{is} \times Post_t\Big) + \theta_i + \theta_t + \varepsilon_{ist}, \tag{11}$$

where $Post_t$ is an indicator for the post-liberalization period (2003-2022 in the annual panel; 2007 and 2022 in the census panels). In the pooled regression, effects are measured relative to the baseline of 1991 and data from 1970 is excluded. The coefficient β^{TH} captures the average effect of U.S. market access on treated and suitable areas relative to controls, pooling all post-1997 years for which data is available. The dynamic and pooled estimates both capture the effects of the same shock, and simply differ in the level of detail provided. The pooled coefficient, β^{TH}, summarizes the long-run effect that the event-study traces over time.

Mechanism: irrigation investment

The pond series is available only within Michoacan at present. I therefore estimate a within-state difference-in-differences at the EA level:

$$Pond_{ist} = \alpha + \sum_{h \neq h_0} \beta_h^H \Big(H_{is} \times \mathbb{1}[t = h] \Big) + \theta_i + \theta_t + \varepsilon_{ist}, \tag{12}$$

with EA fixed effects θ_i, year fixed effects θ_t. Identification requires parallel trends between high- and low-suitability EAs within Michoacan. This would be violated if suitable areas were to experience faster expansion of ponds even in a counterfactual in which the avocado ban lift had never been granted to Michoacan. This may arise, for example, if these areas, by virtue of being more fertile for all types of crops, would have attracted more irrigation investment regardless of avocado demand.

To address this concern, I show parallel trends pre-1997 between suitable and unsuitable areas. This indicates that, prior to the trade shock of interest, outcomes in treated and control areas were evolving similarly, so it is plausible that absent the shock such parallel trends would have continued. Once remote sensing is extended to neighbors, I will move to the full triple difference strategy of 10.

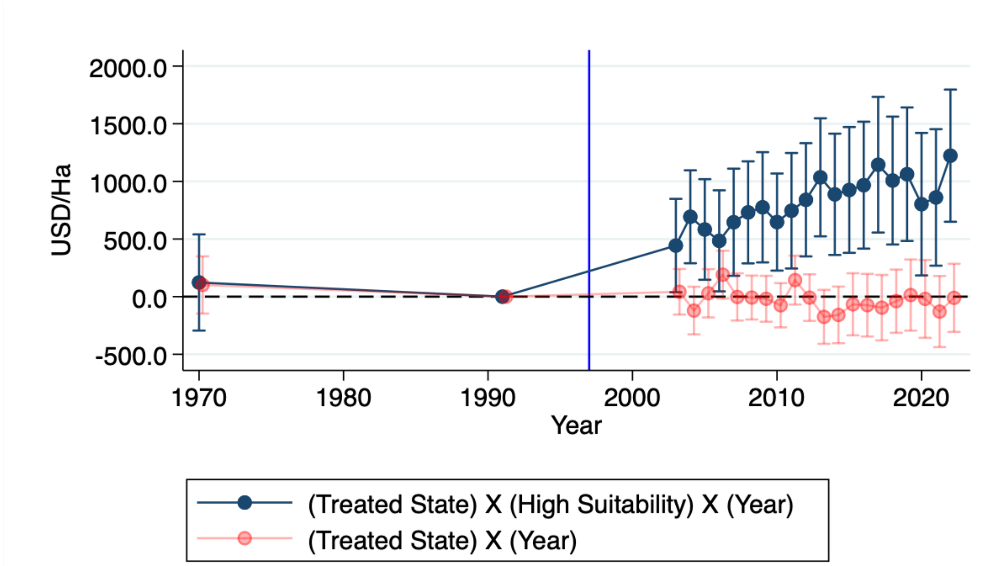
6 Results

In this section, I first report the aggregate productivity effect of the U.S. avocado market opening, and then show its effect on farms' total planted land, crop mix, and physical yields. I then turn to mechanisms, reporting effects on land consolidation and remotely sensed private irrigation investment. I conclude with heterogeneity by land market frictions and specification checks.

6.1 Aggregate output value per hectare (constant prices)

Aggregate value per hectare, based on international 1996 prices, rises sharply in treated areas after 1997 and throughout the 1997-2007 expansion. My main result is contained in Figure 7, which shows that the effects of the trade shock on treated municipalities' output per hectare are on the order of US$ 1,000 per hectare. Baseline (1991) yields are roughly US$ 1,600/ha, so the estimates amount to a 50-60% increase. The effects ramp up in the mid-2000s and persist through 2022.

Figure 7: Triple difference estimates: aggregate value per hectare

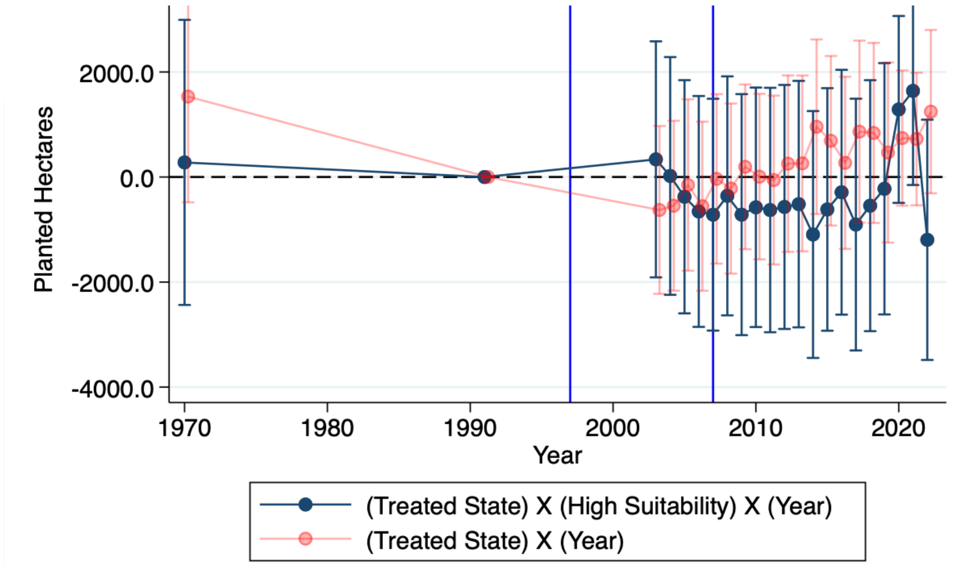


Notes: Estimates of the effects of the avocado ban lift based on equation (10) at the municipality level. The outcome of interest is the aggregate value of agricultural production per hectare (in constant 1996 prices). The coefficients of interest, shown in blue, are the effects of the ban lift on "treated" municipalities, defined as those that are both in the treated state and are highly suitable for avocado cultivation. The coefficients in red capture state-level shocks in Michoacan, common to suitable and unsuitable areas, and are presented to facilitate the interpretation of the triple difference estimates. The blue line in 1997 marks the year in which Michoacan receives the ban lift. All coefficients are plotted alongside their 95% confidence intervals, constructed with standard errors clustered at the municipality level.

As shown in Figure 8, total planted area does not increase differentially in treated-suitable locations, ruling out frontier expansion as the driver of higher values per hectare. This does not mean that the agricultural frontier of the study area is constant, but rather reflects parallel expansions in planted areas across treated and control areas.

Turning to crop-level data, two key findings emerge. First, as shown in Figure 9, land reallocates from maize to avocado in treated areas. Given prevailing physical yields and fixed pre-shock prices, moving planted land from maize to avocado (5-10× value/ha) is sufficient to account for the aggregate gains. Second, Figure 10 shows that physical yields (tons/ha) do not increase within specific crops. For avocado and maize, effects are flat to slightly downwards, suggesting that if anything the shock had a negative impact on physical yields. This is consistent with the entry of young orchards, which decrease physical yield during the adjustment period, and expansion onto less suitable plots. Figures C3 and C4 in the Appendix provide suggestive evidence on these

Figure 8: Triple difference estimates: planted area



Notes: The notes of Figure 7 apply. The outcome of interest is the total planted area of each municipality measured in hectares.

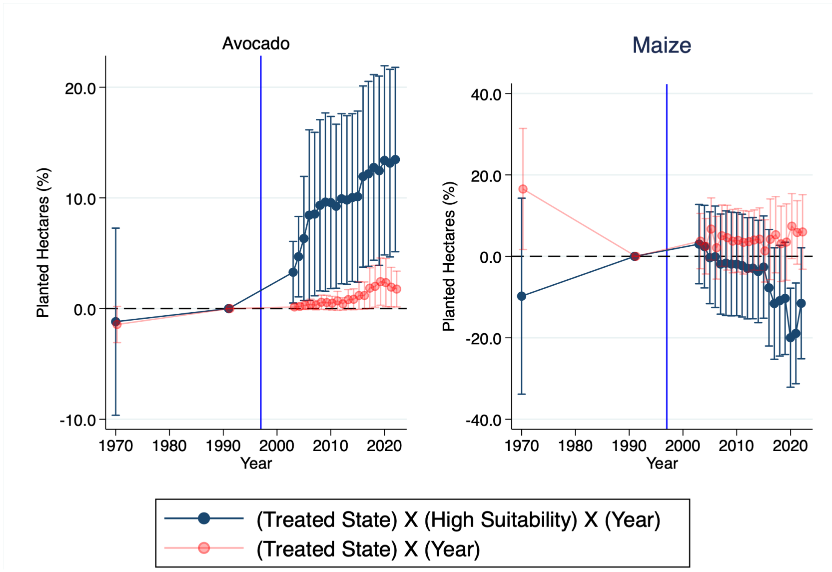
Table 4: Main Results: Triple Difference Estimates

	Agg. Yield	Hectares	Avo. Share	Maize Share	Avo. Yield	Maize Yield
$Post\_t \times T\_s$	-383.94**	1,114.73**	1.07	6.96**	1.61	-2.95***
	(160.52)	(497.27)	(0.77)	(3.21)	(1.07)	(0.52)
$Post\_t \times H\_is$	-453.54***	-167.07	0.35	6.92***	3.63***	-2.46***
	(142.37)	(500.85)	(0.39)	(2.18)	(1.37)	(0.52)
$Post\_t \times H\_is \times T\_s$	1,081.53***	961.06	15.95***	-16.74***	-2.16	1.46**
	(307.50)	(701.53)	(2.11)	(4.37)	(1.58)	(0.58)
Mean. Dep. Var.	1,597.62	13,257.49	2.24	70.48	5.31	4.66
Observations	10,753	10,756	10,753	10,753	4,029	10,746
Adj. R^2	0.82	0.94	0.90	0.89	0.51	0.78

Notes: Each column reports results from a separate regression where the dependent variable is indicated in the column header. *Agg. Yield:* aggregate value of output, at 1996 world prices, divided by total planted area, where both output and area are aggregated across crops. *Hectares:* total planted area across all crops. *Avo. Share:* share of land planted with avocados (%). *Maize Share:* share of land planted with maize (%). *Avo. Yield:* avocado output in metric tons divided by area planted with avocados. The key regressor is the triple interaction between $Post\_t$, an indicator for the post-liberalization period (2003-2022), H_is, an indicator for above-median avocado suitability in municipality i in state s, and T_s, an indicator for the treated state of Michoacan. The data covers 1970, 1991, and 2003-2022. All regressions include municipality and year fixed effects, and are based on the pooled triple differences specification described in presented in equation 11. Standard errors clustered at the municipality level are reported in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

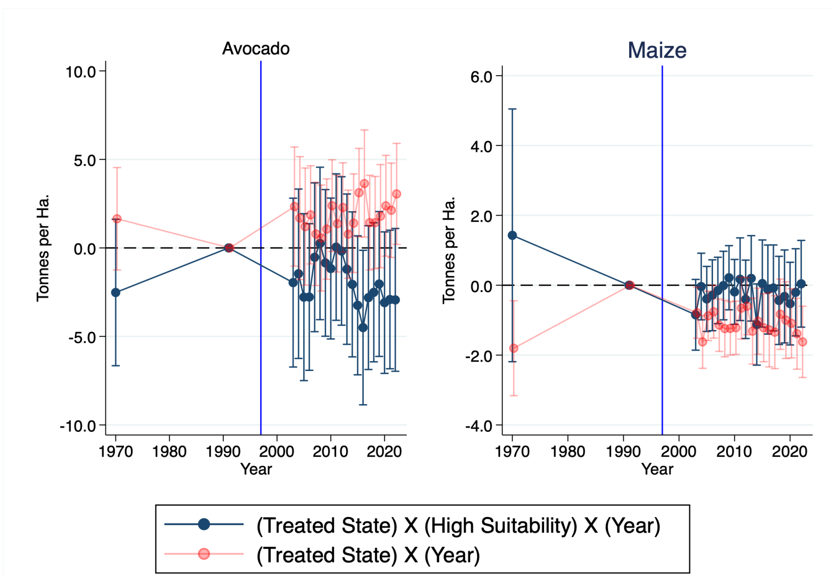
explanations showing that avocado orchards have much lower yields when young and that, in the period after the trade shock, the avocado industry expands into less suitable areas compared to the areas in which avocado was grown prior to 1997. Finally, Table 4 summarizes pooled triple difference estimates for these outcomes, confirming the patterns in the event-study figures.

Figure 9: Triple differences estimates: crop mix



Notes: The notes of Figure 7 apply. The outcomes of interest are the shares of municipalities' planted area devoted to avocado (left panel) and maize (right panel).

Figure 10: Triple differences estimates: physical yields by crop



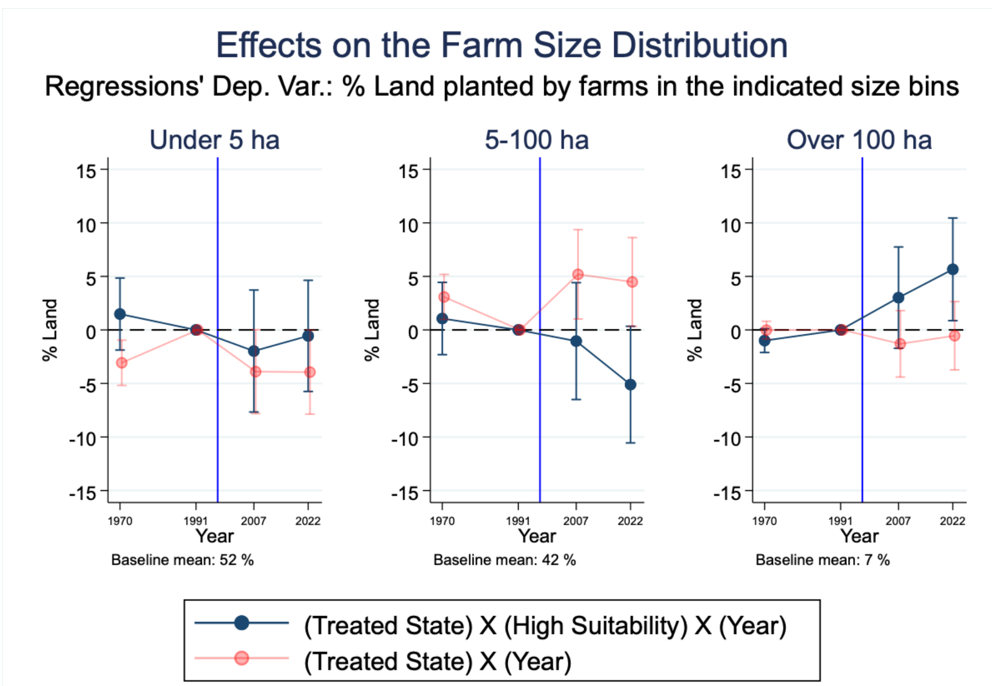
Notes: The notes of Figure 7 apply. The outcomes of interest are municipalities' aggregate physical yield of avocado (left panel) and maize (right panel), in metric tons per planted hectare.

6.2 Mechanisms: Land consolidation and Private Investment

Census microdata show reallocation toward larger farms. As shown in Figure 11, the share of planted land on farms $\geq$ 100 ha approximately doubles in treated areas. Together with the previous result of no response along the extensive margin (i.e., no differential increase in total planted area in Figure 8), this indicates that land reallocation from smaller to larger units is a key margin of adjustment to the trade shock.

This consolidation is driven, specifically, by avocado farms expanding in treated areas. To establish this, as a placebo test, I examine land consolidation among farms that do not cultivate avocados. Figure C5 in the Appendix displays the event-study estimates for the share of land on non-avocado farms larger than 100 hectares. I find no differential expansion of the share of land owned by such farms in treated areas. The 95% confidence intervals on the triple difference coefficients allow us to rule out effects on the order of the main consolidation estimates in Figure 11 (5-10 percentage points). This confirms that the observed land consolidation of Figure 11 is directly attributable to avocado farms, consistent with access to the U.S. avocado market being the underlying driver of land consolidation.

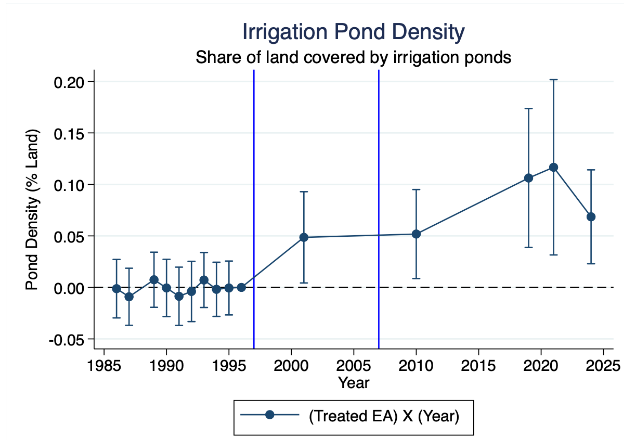
Figure 11: Triple difference estimates: share of land on farms ≥ 100 ha



Notes: The notes of Figure 7 apply. The outcomes of interest are the shares of municipalities' planted area owned by farms in each of three different size categories: less than 5 hectares (left panel), between 5 and 100 hectares (middle panel), and greater than or equal to 100 hectares (right panel).

Descriptives are consistent with the mechanism. As shown in Figures C2, value per hectare rises with size, and avocado cultivation is concentrated among large farms. To connect consolidation to lumpy fixed costs, I measure the construction of on-farm irrigation ponds using a two-stage remote sensing algorithm (Sentinel-2 detection, Landsat backdating). Within Michoacan, a difference-in-differences that compares suitable to unsuitable areas shows a clear post-1997 increase in pond counts/density with parallel pre-trends. Back-of-the-envelope magnitudes imply roughly 20,164 new ponds and, using the average pond costs of Table B1 in Appendix B, about US$ 1.3 billion or nearly 58% of Michoacan's 1996 agricultural GDP in new private irrigation investment between 1997 and 2024.

Figure 12: Difference-in-differences: irrigation ponds (investment)



Notes: Difference-in-differences estimates at the EA level within Michoacan, based on equation (12), comparing high- versus low-suitability EAs. The outcome of interest is the share of EA surface identified as containing irrigation ponds, based on remote sensing. The two blue lines, in 1997 and 2007, mark the beginning and end of the gradual implementation of the U.S. avocado ban lift for Michoacan. All coefficients are plotted alongside their 95% confidence intervals, constructed with standard errors clustered at the EA level.

6.3 Direct evidence on land consolidation from linked census waves

Having presented aggregate evidence that treated areas experience large gains in value per hectare driven by consolidation, I now test directly—at the farm level—whether individual units in treated areas consolidate land. Using a linked Census panel (1991–2007-2022), I estimate farm-level long-difference regressions where the outcome is $\Delta \ln \text{Size}_{j,t}$ for each farm j, between waves 1991–2007 ($t = 2007$) and 2007-2022 ($t = 2022$), measured in log planted hectares. Let $S_{j,t-1}$ denote a baseline farm size measure, either as a large farm indicator equal to one if a farm is larger than 100 hectares or measured continuously in log hectares, both capturing farm size in the *initial* period of the window over which $\Delta \ln \text{Size}_{j,t}$ is computed. Let $i(j)$ and $s(j)$ denote, respectively, the municipality and state where farm j is located. As in equation (10), let $T_{s(j)}$ denote a Michoacan (treated state) indicator and $H_{i(j)s(j)}$ an avocado suitability indicator for the municipality where farm j is located. I estimate equations of the form:

$$
\begin{aligned}
\Delta \ln \text{Size}_{j,t} = \alpha + \beta_1 S_{j,t-1} + \beta_2\big(S_{j,t-1} \times T_{s(j)}\big) + \beta_3\big(S_{j,t-1} \times H_{i(j)s(j)}\big) \\
+ \beta_4\big(S_{j,t-1} \times T_{s(j)} \times H_{i(j)s(j)}\big) + \varphi_{i(j)} + \varepsilon_{j,t}
\end{aligned}
\tag{13}
$$

The design is a farm-level triple difference in long differences. Because treatment and suitability only enter interacted with baseline size, the intercept α captures average growth for small farms in unsuitable municipalities outside the treated state and β_1 measures the difference in farm growth between large and small farms in such municipalities. β_2 and β_3 capture how the large-vs-small growth gap changes in unsuitable treated municipalities and in suitable non-treated municipalities, respectively. The key variation comes from the triple interaction term $S_{j,t-1} \times T_{s(j)} \times H_{i(j)s(j)}$, whose coefficient β_4 measures how the large-vs-small growth gap changes in treated and suitable municipalities, net of the corresponding gaps observed (i) in treated but unsuitable places and (ii) in suitable places outside the treated state.

Table 5 shows estimates of equation (13) for the two different definitions of $S_{j,t-1}$. The estimates confirm that initially larger farms grow differentially faster in the treated municipalities (i.e., suitable municipalities in the treated state). The 1991–2007 results in column (1) are transparent to interpret coefficient-by-coefficient. The estimate of the constant α implies that small farms—those below the 100-hectare cutoff—expanded by about 0.29 log points in unsuitable municipalities outside Michoacan.[20] The estimate of β_1 indicates that, in municipalities that are neither treated nor suitable, the large-versus-small growth gap is strongly negative: relative to small farms, large farms grew 0.57 log points less. Evaluated in levels, the implied growth for large farms in that baseline cell is $\alpha + \beta_1 = 0.29 - 0.57 = -0.28$ log points (about a 24% contraction). The estimate for the β_2 coefficient for the $S_{j,t-1} \times T_{s(j)}$ interaction indicates that this gap in farm size growth between small and large farms is not significantly different in Michoacan compared to the control states. In contrast, the estimate for the β_3 coefficient for the $S_{j,t-1} \times H_{i(j)s(j)}$ interaction is negative and significant, and indicates that in suitable municipalities outside Michoacan the large-versus-small growth gap is further exacerbated by roughly 0.10 log points. Finally, the estimate on the coefficient of interest β_4 for the triple interaction term is statistically significant and positive: in treated municipalities of the treated state, large farms grow 0.08 log points faster compared to

[20]In general, because the regressions include municipality fixed effects, interpreting estimates of the constant α requires care. My interpretation relies on the fact that I normalize municipality fixed effects to sum to zero.

suitable non-treated municipalities and compared to unsuitable treated municipalities. Under the identifying assumption of parallel trends, this estimate provides direct evidence that the trade shock triggers land consolidation.

The continuous specification in column (2) confirms this pattern and assuages concerns about the arbitrary 100 hectares cutoff used in column (1) spuriously driving a consolidation result. The later panel (2007-2022) exhibits qualitatively similar patterns, with larger estimated magnitudes for the coefficient of interest β_4 (columns (3) and (4)). These results must be interpreted with caution because the underlying linked panel has a match rate around 40% between waves, and hence is not fully representative of all farms. But this evidence, taken together with the previous subsection, provides compelling evidence of land consolidation as a key consequence of the trade shock this paper studies.

6.4 Effect heterogeneity by land market frictions

I conclude by testing the model's prediction that higher communal land shares (ejidos) dampen consolidation and productivity. I do so by comparing productivity gains, consolidation, and investment, in areas where ejido prevalence is low and in areas where ejidos dominate. I show, in Tables 6 and 7, respectively, that productivity and consolidation effects are concentrated in areas with low ejido prevalence. Investment responses, estimated separately in high- and low-friction areas in Figure 13, mirror this heterogeneity.

Table 6 quantifies this pattern precisely. In areas with low ejido prevalence (below the median share of communal land), the share of agricultural land on large farms ($\geq$ 100 ha) increases by 9.1 percentage points by 2022—a doubling relative to the 1991 baseline. In contrast, high-ejido areas exhibit essentially zero consolidation (point estimate of -0.0002, statistically indistinguishable from zero). This stark divergence directly confirms the mechanism: where land markets function, trade shocks trigger consolidation and productivity gains; where institutional frictions bind, the same export opportunities fail to generate structural transformation. The results demonstrate that market access alone is insufficient—reallocation requires institutions that permit land to move to its highest-value use. Finally, Table 7 confirms that productivity gains follow the same pattern: output per hectare rises by around \$2,000 in low-ejido areas (significant at the 1% level in both 2007 and 2022), while high-ejido areas show statistically and economically insignificant effects. The parallel heterogeneity in both consolidation and productivity validates that land reallocation

Table 5: Baseline Growth and Interaction Effects

	Outcome: $\Delta \ln \text{Size}_i$			
	(1)	(2)	(3)	(4)
	(0.002)	(0.004)	(0.002)	(0.003)
Large Farm (β_1)	-0.5682***		-0.4259***	
	(0.005)		(0.003)	
Large Farm $\times T_{s(j)}$ (β_2)	-0.0115		-0.1111***	
	(0.013)		(0.011)	
Large Farm $\times H_{i(j)s(j)}$ (β_3)	-0.0954***		-0.1826***	
	(0.007)		(0.006)	
Large Farm $\times T_{s(j)} \times H_{i(j)s(j)}$ (β_4)	0.0771***		0.1391***	
	(0.017)		(0.015)	
Log Size (Ha.) (β_1)		-0.3434***		-0.5520***
		(0.003)		(0.004)
Log Size (Ha.) $\times T_{s(j)}$ (β_2)		-0.2047***		-0.0354***
		(0.011)		(0.010)
Log Size (Ha.) $\times H_{i(j)s(j)}$ (β_3)		-0.1643***		-0.1499***
		(0.006)		(0.006)
Log Size (Ha.) $\times T_{s(j)} \times H_{i(j)s(j)}$ (β_4)		0.1593***		0.1092***
		(0.015)		(0.013)
Constant (α)	0.2919***	0.8012***	0.1260***	0.7158***
Observations	205,547	205,547	287,847	287,847
R^2	0.186	0.303	0.134	0.335
Fixed Effects	Municipality	Municipality	Municipality	Municipality
Period	1991-2007	1991-2007	2007-2022	2007-2022

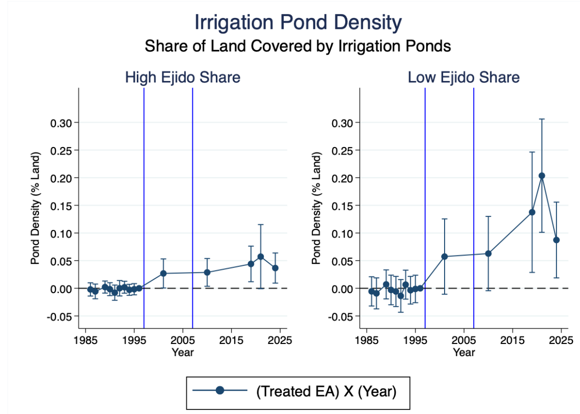
Notes: Standard errors in parentheses. Farm sizes are measured hectares of planted land aggregated across crops. *Large Farm* is an indicator for farms above 100 hectares. $T_{s(j)}$ is an indicator for treatment state (Michoacan). $H_{i(j)s(j)}$ is an indicator for high suitability (above median within Michoacan). All regressions include municipality fixed effects and are based on equation 13. Outcomes are measured as the change in the natural logarithm of farm size between agricultural census years. Columns (1) and (2) use data from the 1991 and 2007 agricultural censuses; columns (3) and (4) use data from the 2007 and 2022 agricultural censuses. *** p<0.01, ** p<0.05, * p<0.1.

is the proximate driver of aggregate gains.

7 Conclusion

This paper shows that opening rich-country markets can trigger large, efficiency-enhancing reallocation in developing-country agriculture. Exploiting the 1997 removal of U.S. phytosanitary barriers on Mexican avocados, granted access only to the state of Michoacan, I find sizable increases in aggregate agricultural productivity. In treated areas, value per hectare rises by over \$1000/ha, driven primarily by a shift of land from maize to avocado—a crop nearly ten times more valuable per hectare. This adjustment operates through land consolidation, with the share of land culti-

Figure 13: Difference-in-differences estimates of investment effects, by baseline ejido prevalence.



Notes: The notes of Figure 12 apply. The estimates are shown separately for EAs with below-median (left panel) and above-median (right panel) baseline ejido prevalence. This median is computed in 1991 across all EAs in Michoacan.

Table 6: Difference-in-differences estimates of consolidation effects, by baseline ejido prevalence.

	All (1)	**Low Ejido** (2)	**High Ejido** (3)
Dependent variable: % Land Above 100 Ha.			
2007 × Treat × Suitable	-0.0047	-0.0180	0.0161
	(0.025)	(0.036)	(0.034)
2022 × Treat × Suitable	0.0557**	0.0911**	-0.0002
	(0.026)	(0.040)	(0.028)
Observations	7,252	3,803	3,449
Fixed effects	State×Year	State×Year	State×Year
Sample	All	Low Ejido	High Ejido

Note: Low and High ejido areas are defined as those where, in 1991, the share of agricultural land held under communal tenure is below or above the median across all EAs in Michoacan. Treat indicates Michoacan; Suitable indicates high avocado suitability. Standard errors clustered at the municipality level in parentheses. *** $p < 0.01$.

vated by large farms (over 100 hectares) approximately doubling. I document a surge in lumpy irrigation investments (pond construction), confirming that farms incurred substantial fixed costs to enter the global avocado trade. The presence of these fixed costs is consistent with a theoretical framework in which trade increases the economies of scale in agriculture, helping rationalize the observed productivity and consolidation effects. I also find that land market frictions shape the gains from trade: all effects are concentrated in areas with low collective landholdings at baseline.

My contribution is threefold. First, I provide direct evidence that trade liberalization induces between-farm reallocation that raises aggregate productivity. In doing so, I build a bridge between well-established but largely disconnected literatures on agricultural productivity and the

Table 7: Differences-in-differences estimates of yield, by baseline ejido prevalence.

	All (1)	**Low Ejido** (2)	**High Ejido** (3)
Dependent variable: Aggregate Value of Output per Ha. (1996 world prices)			
2007 × Treat × Suitable	1145.79**	1979.56**	-240.77
	(549.38)	(769.93)	(321.51)
2022 × Treat × Suitable	1612.23***	2658.67***	-1.20
	(602.05)	(799.80)	(464.39)
Observations	7,252	3,803	3,449
Fixed effects	State×Year	State×Year	State×Year
Sample	All	Low Ejido	High Ejido

Note: Low and High ejido areas are defined as those where, in 1991, the share of agricultural land held under communal tenure is below or above the median across all EAs in Michoacan. Treat indicates Michoacan; Suitable indicates high avocado suitability. Standard errors clustered at the municipality level in parentheses. *** $p < 0.01$.

productivity consequences of trade. While much prior empirical work on trade has focused on within-firm performance gains in manufacturing, my results show that, in agriculture, a central margin is the reallocation of land toward large farms that can afford lumpy investments. Second, building on this evidence, I use a novel methodology to measure lumpy investments in agriculture, allowing me to provide direct evidence on the role of fixed costs emphasized by trade theory: export opportunities raise returns to scale which in turn leads to large-firm expansion. In my setting, large farms are the ones able to finance indivisible irrigation infrastructure required to plant avocado orchards, and the trade shock triggers land consolidation by boosting these investments' returns. Third, I show that the potentially large gains from trade suggested by the main results are not automatic but instead depend on local institutions. In particular, land market frictions arising from collective ownership can dampen most of the gains. I show that, where collective land tenure is prevalent and land transactions are costly, consolidation and productivity growth are muted, consistent with an active literature on how land market frictions may cause costly distortions in developing country agriculture (De Janvry et al., 2015; Manysheva, 2022).

A natural next step in this research is to assess whether the mechanisms identified here extend beyond Mexico. In sub-Saharan Africa, for instance, relaxing European sanitary standards might trigger a similar move toward high-value export crops, provided land can be consolidated and fixed costs financed. The results suggest that, if such reforms occur, productivity gains will be concentrated in areas where land markets facilitate reallocation toward large farms. Where land is fragmented, customary tenure dominant, or transaction costs high, the supply response will be dampened. Thus, external liberalization and internal market institutions jointly determine the realized gains from trade.

Some limitations qualify these results and leave open questions for future work. First, the environmental consequences of the trade boom studied in this paper deserve careful study. Avocado cultivation is water-intensive and orchard expansion has been linked to localized water stress and deforestation in public discussion. A full welfare assessment of the trade shock studied by this paper should incorporate these external costs. Second, the distributional incidence of the boom is ambiguous ex ante. Even as aggregate productivity rises, the gains may accrue disproportionately to larger landowners who can consolidate and invest, with smallholders benefiting only indirectly (e.g., through employment or land sales). Third, measuring general equilibrium effects lies beyond the scope of this paper. The expansion of Michoacan's avocado sector may, for instance, affect the returns to avocado cultivation in control states. The shock studied here led to avocado prices roughly doubling throughout Mexico, and this in turn created incentives for other states to switch from maize to avocados just as in Michoacan. The empirical strategy, as with most studies of this nature, relies on local variation in treatment exposure to identify causal effects, and implicitly assumes away nationwide effects of the shock which may affect both treated and control areas. Finally, the remote-sensing measure of investment, while validated, will be further refined in future work with recently acquired proprietary imagery, and it will be extended to all of Mexico beyond just Michoacan.

Taken together, these findings have two key policy implications. The first one is that developed countries can contribute to agricultural productivity growth abroad by systematically reducing non-tariff barriers on high-value crops, where sanitary and phytosanitary restrictions remain widespread. The second is that trade liberalization alone may not be sufficient. Where communal land tenure prevents consolidation, developing countries undergoing trade liberalization episodes should consider reforms that allow land to be leased or sold more easily (such as Mexico's 1992 PROCEDE program, but more aggressively implemented). Policy evaluations that take into consideration interactions with international trade liberalization remain a fruitful avenue for future research. The broader lesson is that trade policy may work best when complemented by reforms that reduce frictions in the markets through which adjustment to trade occurs.

References

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. (2022). Geography and agricultural productivity: Cross-country evidence from micro plot-level data. *The Review of Economic Studies*, 89(4):1629–1653.

Atkin, D. and Khandelwal, A. K. (2020). How distortions alter the impacts of international trade in developing countries. *Annual Review of Economics*, 12(1):213–238.

Atkin, D., Khandelwal, A. K., and Osman, A. (2017). Exporting and firm performance: Evidence from a randomized experiment. *The quarterly journal of economics*, 132(2):551–615.

Bai, Y., Jin, K., and Lu, D. (2024). Misallocation under trade liberalization. *American Economic Review*, 114(7):1949–1985.

Beghin, J. C., Maertens, M., and Swinnen, J. (2015). Non-Tariff measures and standards in trade and global value chains. *Annual Review of Resource Economics*. Forthcoming; see Bioeconomics Working Paper 2015/02 for the preprint.

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2007). Firms in international trade. *Journal of Economic perspectives*, 21(3):105–130.

Bredahl, M. E. (2001). Trade Liberalization Under NAFTA: Trade in Avocados.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on Argentinian firms. *American economic review*, 101(1):304–340.

Chen, C., Restuccia, D., and Santaeulàlia-Llopis, R. (2023). Land misallocation and productivity. *American Economic Journal: Macroeconomics*, 15(2):441–465.

Costinot, A. and Donaldson, D. (2016). How large are the gains from economic integration? Theory and evidence from US agriculture, 1880-1997. Technical report, National Bureau of Economic Research.

Couttenier, M., Di Rollo, S., Inguere, L., Mohand, M., and Schmidt, L. (2022). Mapping artisanal and small-scale mines at large scale from space with deep learning. *Plos one*, 17(9):e0267963.

De Haro, I. (2022). Avocados: Mexico's green gold. The US opioid crisis and its impact on Mexico's drug cartel violence. Working Paper.

De Janvry, A., Dequiedt, V., and Sadoulet, E. (2014). The demand for insurance against common shocks. *Journal of Development Economics*, 106:227–238.

De Janvry, A., Emerick, K., Gonzalez-Navarro, M., and Sadoulet, E. (2015). Delinking land rights from land use: Certification and migration in Mexico. *American Economic Review*, 105(10):3125–3149.

Eastwood, R., Lipton, M., and Newell, A. (2010). Farm size. *Handbook of agricultural economics*, 4:3323–3397.

FAO (2016). ECOCROP database. Food and Agriculture Organization of the United Nations.

Farrokhi, F. and Pellegrina, H. S. (2023). Trade, technology, and agricultural productivity. *Journal of Political Economy*, 131(9):2509–2555.

Felix, M. (2021). Trade, labor market concentration, and wages. *Job Market Paper*, 64.

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.

Gaigne, C. and Gouel, C. (2022). Trade in agricultural and food products. In *Handbook of Agricultural Economics*, volume 6, pages 4845–4931.

Goldberg, P. K. and Reed, T. (2023). Demand-side constraints in development. the role of market size, trade, and (in) equality. *Econometrica*, 91(6):1915–1950.

Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of political Economy*, 116(6):981–1022.

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

Gourdon, J., Cadot, O., Ghersi, G., and Vanzetti, D. (2015). Estimating the price effects of non-tariff measures. *World Trade Review*, 14(4):681–703.

Grundke, R. and Moser, C. (2019). Hidden protectionism? Evidence from non-tariff barriers to trade in the United States. *Journal of International Economics*, 117:143–157.

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

Hopenhayn, H. A. (2014). Firms, misallocation, and aggregate productivity: A review. *Annu. Rev. Econ.*, 6(1):735–770.

Ignatenko, A., Lashkaripour, A., Macedoni, L., and Simonovska, I. (2025). Making america great again? the economic impacts of liberation day tariffs. *Journal of International Economics*. Forthcoming.

Khachiyan, A., Thomas, A., Zhou, H., Hanson, G., Cloninger, A., Rosing, T., and Khandelwal, A. K. (2022). Using neural networks to predict microspatial economic growth. *American Economic Review: Insights*, 4(4):491–506.

Lagakos, D. and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2):948–980.

Lowder, S. K., Sánchez, M. V., and Bertini, R. (2021). Which farms feed the world and has farmland become more concentrated? *World Development*, 142:105455.

Lowder, S. K., Skoet, J., and Raney, T. (2016). The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World development*, 87:16–29.

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

McCaig, B. and Pavcnik, N. (2018). Export markets and labor allocation in a low-income country. *American Economic Review*, 108(7):1899–1941.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Melitz, M. J. and Redding, S. J. (2015). New trade models, new welfare implications. *American Economic Review*, 105(3):1105–1146.

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants. *The Review of Economic Studies*, 69(1):245–276.

Rao, E., Brümmer, B., and Qaim, M. (2012). Farmer participation in supermarket channels, production technology, and efficiency: The case of vegetables in kenya. *American Journal of Agricultural Economics*, 94(4):891–912.

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.

SAGARPA (2018). Lineamientos y criterios para el ejercicio de recursos en las entidades federativas en materia de información (snidrus 2018). Technical report, Servicio de Información Agroalimentaria y Pesquera (SIAP).

Segerstrom, P. S. and Sugita, Y. (2015). The impact of trade liberalization on industrial productivity. *Journal of the European Economic Association*, 13(6):1167–1179.

Sotelo, S. (2020). Domestic trade frictions and agriculture. *Journal of Political Economy*, 128(7):2690–2738.

Stanford, L. (2002). Constructing"quality": The political economy of standards in Mexico's avocado industry. *Agriculture and Human Values*, 19(4):293–310.

Suri, T. and Udry, C. (2022). Agricultural technology in Africa. *Journal of Economic Perspectives*, 36(1):33–56.

Trefler, D. (2004). The long and short of the canada-us free trade agreement. *American Economic Review*, 94(4):870–895.

Wollburg, P., Bentze, T., Lu, Y., Udry, C., and Gollin, D. (2024). Agricultural productivity growth in Africa: New evidence from microdata. *Proceedings of the National Academy of Sciences*.

A Appendix to Theoretical Framework

A.1 Derivation of Aggregate Output under Autarky

Farm-level output is $y(\varphi) = \varphi\ell(\varphi)^{\alpha}$. I substitute farm-level land demand from equation (1), $\ell_B(\varphi) = \left(\frac{\alpha\varphi}{w}\right)^{\gamma}$, into the previous equation, which yields $y(\varphi) = \left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}\varphi^{\gamma}$. Next, relying on the index $\widetilde{\varphi} = \left(\int_{\varphi_{min}}^{\infty}\varphi^{\gamma}g(\varphi)d\varphi\right)^{1/\gamma}$ and integrating over all farms, I obtain aggregate output:

$$Y = \int_{\varphi_{min}}^{\infty} y(\varphi)g(\varphi)\,d\varphi = \widetilde{\varphi}^{\gamma}\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}.$$

Then, I rearrange equation (2) to obtain $\frac{\alpha}{w} = \widetilde{\varphi}^{-1}L^{1-\alpha}$. Substituting this expression into the previous equation for aggregate output, I obtain:

$$Y = \widetilde{\varphi}^{\gamma}\left(\widetilde{\varphi}^{-1}L^{1-\alpha}\right)^{\frac{\alpha}{1-\alpha}} = \widetilde{\varphi}L^{\alpha}.$$

A.2 Derivation of the Price of Land in the Open Economy

Starting from the market-clearing condition under trade, I have:

$$L = \int_{\varphi^*}^{\infty}\ell_A(\varphi)g(\varphi)d\varphi + \int_{\varphi_{min}}^{\varphi^*}\ell_B(\varphi)g(\varphi)d\varphi.$$

Replacing the land demand expressions from equations (4) and (1), I obtain:

$$\begin{aligned} L &= \left(\frac{\alpha}{w}\right)^{\gamma}\rho\int_{\varphi^*}^{\infty}\varphi^{\gamma}g(\varphi)d\varphi + \left(\frac{\alpha}{w}\right)^{\gamma}\int_{\varphi_{min}}^{\varphi^*}\varphi^{\gamma}g(\varphi)d\varphi \\ &= \left(\frac{\alpha}{w}\right)^{\gamma}\left[(\rho-1)\int_{\varphi^*}^{\infty}\varphi^{\gamma}g(\varphi)d\varphi + \int_{\varphi_{min}}^{\infty}\varphi^{\gamma}g(\varphi)d\varphi\right] \\ &= \left(\frac{\alpha}{w}\right)^{\gamma}\widetilde{\varphi}^{\gamma}\left[(\rho-1)R+1\right] \end{aligned}$$

And rearranging the last expression to solve for w, I obtain the equilibrium price of land under trade:

$$w_T = \alpha\widetilde{\varphi}[1+(\rho-1)R]^{1-\alpha}L^{-(1-\alpha)}.$$

A.3 Derivation of a Closed Form for $C(\lambda)$

Using (4) and (1), I can rewrite $C(\lambda)$ as

$$C(\lambda) = \left(\frac{\alpha}{w_T}\right)^{\gamma} \rho \int_{\varphi_p^*}^{\infty} \varphi^{\gamma} g(\varphi) d\varphi - \left(\frac{\alpha}{w_0}\right)^{\gamma} \int_{\varphi_p^*}^{\infty} \varphi^{\gamma} g(\varphi) d\varphi. \tag{A.3.1}$$

For private farms, it is useful to define the share of land owned under autarky by those who switch to sector A once the economy opens to trade as

$$R(\lambda) \equiv \frac{\int_{\varphi_p^*}^{\infty} \varphi^{\gamma} g(\varphi) d\varphi}{\int_{\varphi_{min}}^{\infty} \varphi^{\gamma} g(\varphi) d\varphi}. \tag{A.3.2}$$

Using equation (A.3.2), I can rewrite (A.3.1) as:

$$C(\lambda) = (\alpha\widetilde{\varphi})^{\gamma} R(\lambda) \left[\frac{\rho}{w_T^{\gamma}} - \frac{1}{w_0^{\gamma}}\right] \tag{A.3.3}$$

Next, I use the land prices w_T and w_0, which in the economy with frictions follow a modified version of equations (6) and (3) to reflect that the total supply of tradable land is $(1-\lambda)L$ instead of L:

$$\begin{aligned} w_T &= \alpha\widetilde{\varphi}\left[1 + (\rho - 1)R(\lambda)\right]^{1-\alpha} ((1-\lambda)L)^{-(1-\alpha)} \\ w_0 &= \alpha\widetilde{\varphi}((1-\lambda)L)^{-(1-\alpha)} \end{aligned}$$

Replacing the expressions for land prices into (A.3.3) leads, after some algebraic manipulation, to the following closed form for $C(\lambda)$:

$$C(\lambda) = (1-\lambda)LR(\lambda)\left[\frac{(\rho - 1)(1 - R(\lambda))}{1 + (\rho - 1)R(\lambda)}\right] \tag{A.3.4}$$

A.4 Proof of Proposition 3

First, I establish that $R(\lambda)$ is decreasing in λ. This can be shown by contradiction. Take $\lambda_2 > \lambda_1$ and suppose $R(\lambda_2) > R(\lambda_1)$. Since the total supply of tradable land decreases with λ, then w_T must increase. A higher land price w_T increases the cutoff productivity φ_p^*, reducing the set of farms that switch to sector A. But then, this would imply $R(\lambda_2) < R(\lambda_1)$.

Next I compute the derivative of $C(\lambda)$ with respect to λ. It is convenient to rewrite (A.3.4) as:

$$C(\lambda) = (1-\lambda)L(\rho-1)f(R(\lambda)),$$

where $f(R) \equiv \frac{R(1-R)}{1+(\rho-1)R}$ with derivative $f'(R) = \frac{1-2R-(\rho-1)R^2}{(1+(\rho-1)R)^2}$. The derivative of $C(\lambda)$ is

$$C'(\lambda) = -L(\rho-1)f(R(\lambda)) + (1-\lambda)L(\rho-1)f'(R(\lambda))R'(\lambda) \tag{A.4.1}$$

Since $R'(\lambda) < 0$ (established at the start of this proof), I can write $R'(\lambda) = -|R'(\lambda)|$, so:

$$C'(\lambda) = -L(\rho-1)\Big[f(R) + (1-\lambda)f'(R)|R'(\lambda)|\Big]. \tag{A.4.2}$$

If $f'(R) > 0$, then the term in brackets in the last equation is positive and $C'(\lambda) < 0$ follows. To complete the proof, I need to establish that

$$f(R) > (1-\lambda)|f'(R)|\,|R'(\lambda)|. \tag{A.4.3}$$

From the private-land price (displayed above),

$$w_T = \alpha\widetilde{\varphi}\,[1+(\rho-1)R(\lambda)]^{1-\alpha}\Big((1-\lambda)L\Big)^{-(1-\alpha)}.$$

Taking logs and differentiating w.r.t. λ gives

$$\frac{d\ln w_T}{d\lambda} = (1-\alpha)\left[\frac{1}{1-\lambda} + \frac{(\rho-1)R'(\lambda)}{1+(\rho-1)R(\lambda)}\right] > 0,$$

where the inequality follows from the fact that $R'(\lambda)$ is negative.[21] Rearranging this last inequality yields $|R'(\lambda)| \leq \frac{1+(\rho-1)R(\lambda)}{(\rho-1)(1-\lambda)}$, which I plug into (A.4.3):

$$f(R) > |f'(R)|\,\frac{1+(\rho-1)R}{(\rho-1)}. \tag{A.4.4}$$

[21] The derivative of R w.r.t. λ has the opposite sign to the derivative of w_T w.r.t the same parameter. This follows from the fact that R is a decreasing function of φ_p^*, which is in turn an increasing function of w_T. There is no other channel through which λ affects R.

After some manipulation,[22] I obtain the condition:

$$h(R) \equiv 1 + R(\rho - 3) - 2(\rho - 1)R^2 > 0. \tag{A.4.5}$$

The LHS of this inequality is a concave quadratic function in R with roots

$$R_1 = \frac{\rho - 3 + \sqrt{(\rho-3)^2 + 8(\rho-1)}}{4(\rho-1)} = \frac{1}{2},$$
$$R_2 = \frac{\rho - 3 - \sqrt{(\rho-3)^2 + 8(\rho-1)}}{4(\rho-1)} = -\frac{1}{\rho-1}.$$

Because $h(R)$ is concave, it is positive whenever R lies between R_1 and R_2. As established before, $R(\lambda)$ is decreasing in λ, so the condition that $R(\lambda) < \frac{1}{2}$ for all λ ensures that $h(R(\lambda))$ is positive for all λ. □

A.5 Proof of Proposition 4

Aggregate output is the sum of private and collective components. Under autarky, the private sector's output is $Y_{0,p} = \widetilde{\varphi}((1-\lambda)L)^{\alpha}$. In the collective sector each plot's output is independent of λ because their size $\ell_c = \lambda L/N_c$ is fixed. Total output for the collective sector is $Y_{0,c} = N_c \ell_c^{\alpha} \int_{\varphi_{min}}^{\infty} \varphi g(\varphi)\, d\varphi = \lambda L \ell_c^{\alpha-1} \int_{\varphi_{min}}^{\infty} \varphi g(\varphi)\, d\varphi$. Under trade, $Y_{T,p}(\lambda) = \widetilde{\varphi}((1-\lambda)L)^{\alpha}[1 + (\rho-1)R(\lambda)]^{1-\alpha}$, and $Y_{T,c}(\lambda) = \lambda L \ell_c^{\alpha-1}\left(\int_{\varphi_{min}}^{\infty} \varphi g(\varphi)\, d\varphi + (p_A - 1)\int_{\varphi_c^*}^{\infty} \varphi g(\varphi)\, d\varphi\right)$. Therefore, the total productivity gains from trade are

$$\Delta Y(\lambda) = \underbrace{\widetilde{\varphi}\Big((1-\lambda)L\Big)^{\alpha}\Big([1 + (\rho-1)R(\lambda)]^{1-\alpha} - 1\Big)}_{\equiv \Delta Y_p(\lambda)} + \underbrace{\lambda L \ell_c^{\alpha-1}(p_A - 1)\int_{\varphi_c^*}^{\infty} \varphi\, g(\varphi)\, d\varphi}_{\equiv \Delta Y_c(\lambda)}.$$

Differentiating with respect to λ:

[22] Replace $f(R)$ and $f'(R)$ with their expressions to obtain the inequality $\frac{R(1-R)}{1+(\rho-1)R} > \frac{|1-2R-(\rho-1)R^2|}{(1+(\rho-1)R)^2} \cdot \frac{1+(\rho-1)R}{(\rho-1)}$, which simplifies into $R(1-R)(\rho-1) > |1 - 2R - (\rho-1)R^2| = (\rho-1)R^2 + 2R - 1$, with the last equality following from the fact that I am considering the case with $f'(R) < 0$.

$$\frac{d\Delta Y_p}{d\lambda} = \widetilde{\varphi} L^{\alpha}\Big[-\alpha(1-\lambda)^{\alpha-1}\Big([1+(\rho-1)R(\lambda)]^{1-\alpha}-1\Big) + (1-\lambda)^{\alpha}(1-\alpha)[1+(\rho-1)R(\lambda)]^{-\alpha}(\rho-1)R'(\lambda)\Big],$$

$$\frac{d\Delta Y_c}{d\lambda} = L\ell_c^{\alpha-1}(p_A-1)\int_{\varphi_c^*}^{\infty}\varphi g(\varphi)\,d\varphi.$$

As established in the proof of Proposition 3, $R'(\lambda) < 0$. Since $\rho > 1$ and $R(\lambda) \geq 0$, the bracketed term $[1+(\rho-1)R]^{1-\alpha} - 1$ is strictly positive. With $R'(\lambda) < 0$, the second term in $\frac{d\Delta Y_p}{d\lambda}$ is nonpositive and dropping it yields a valid lower bound on the magnitude of private losses. I must verify that these private losses dominate collective gains for all $\lambda \in [0,1)$.

I begin by establishing a bound on tail integrals. For any $t > 0$, I claim that

$$\int_t^{\infty}\varphi g(\varphi)\,d\varphi \leq \frac{\mathbb{E}[\varphi^2]}{t}.$$

To see this, note that for any $\varphi \geq t > 0$, we have $\varphi = \varphi \cdot 1 \leq \varphi \cdot \frac{\varphi}{t} = \frac{\varphi^2}{t}$. Multiplying both sides by $g(\varphi)$ and integrating from t to ∞ yields

$$\int_t^{\infty}\varphi g(\varphi)\,d\varphi \leq \int_t^{\infty}\frac{\varphi^2}{t}g(\varphi)\,d\varphi = \frac{1}{t}\int_t^{\infty}\varphi^2 g(\varphi)\,d\varphi \leq \frac{1}{t}\int_{\varphi_{min}}^{\infty}\varphi^2 g(\varphi)\,d\varphi = \frac{\mathbb{E}[\varphi^2]}{t},$$

where the second inequality follows because $\varphi^2 g(\varphi) \geq 0$ everywhere (recall that φ_{min} is the minimum point of the support of $g(\varphi)$).

Applying this bound with $t = \varphi_c^* = \frac{f_A}{(p_A-1)\ell_c^{\alpha}}$ (the collective export cutoff), I obtain

$$\int_{\varphi_c^*}^{\infty}\varphi g(\varphi)\,d\varphi \leq \frac{\mathbb{E}[\varphi^2]}{\varphi_c^*} = \frac{(p_A-1)\ell_c^{\alpha}\mathbb{E}[\varphi^2]}{f_A}.$$

Therefore,

$$\frac{d\Delta Y_c}{d\lambda} = L\ell_c^{\alpha-1}(p_A-1)\int_{\varphi_c^*}^{\infty}\varphi g(\varphi)\,d\varphi \leq L\frac{(p_A-1)^2\mathbb{E}[\varphi^2]}{f_A}\ell_c^{2\alpha-1}.$$

Since $\alpha > \frac{1}{2}$ by assumption (ii), I have $2\alpha - 1 > 0$, so collective gains vanish as $\ell_c \to 0$.

To bound private losses from below, I drop the negative $R'(\lambda)$ term from $\frac{d\Delta Y_p}{d\lambda}$ to obtain

$$\left|\frac{d\Delta Y_p}{d\lambda}\right| \geq \alpha(1-\lambda)^{\alpha-1}\widetilde{\varphi} L^{\alpha}\Big([1+(\rho-1)R(\lambda)]^{1-\alpha}-1\Big).$$

Because the function $(1+x)^{1-\alpha}$ is concave, for $x = (\rho - 1)R(\lambda) \in [0, \rho - 1]$,

$$(1+x)^{1-\alpha} - 1 \;\geq\; (1-\alpha)\rho^{-\alpha}\, x,$$

where $(1 + \rho - 1)^{-\alpha} = \rho^{-\alpha}$. Therefore,

$$\left|\frac{d\Delta Y_p}{d\lambda}\right| \;\geq\; \alpha(1-\alpha)\rho^{-\alpha}(\rho-1)\,\widetilde{\varphi}\,L^{\alpha}\,(1-\lambda)^{\alpha-1}R(\lambda).$$

By assumption (iii), $(1-\lambda)^{\alpha-1}R(\lambda) \geq m$ for all $\lambda \in [0, 1)$, so

$$\left|\frac{d\Delta Y_p}{d\lambda}\right| \;\geq\; \alpha(1-\alpha)\rho^{-\alpha}(\rho-1)\,\widetilde{\varphi}\,L^{\alpha}\,m$$

for all λ.

To prove the proposition, private losses from raising λ dominate collective gains:

$$\alpha(1-\alpha)\rho^{-\alpha}(\rho-1)\,\widetilde{\varphi}\,L^{\alpha}\,m \;>\; L\,\frac{(p_A-1)^2\mathbb{E}[\varphi^2]}{f_A}\,\ell_c^{2\alpha-1}.$$

Rearranging:

$$\ell_c^{2\alpha-1} < \frac{\alpha(1-\alpha)\rho^{-\alpha}(\rho-1)\,\widetilde{\varphi}\,f_A\,L^{\alpha-1}m}{(p_A-1)^2\mathbb{E}[\varphi^2]}.$$

Since $2\alpha - 1 > 0$:

$$\ell_c < \left[\frac{\alpha(1-\alpha)\rho^{-\alpha}(\rho-1)\,\widetilde{\varphi}\,f_A\,m}{(p_A-1)^2\mathbb{E}[\varphi^2]}\right]^{\frac{1}{2\alpha-1}} L^{\frac{\alpha-1}{2\alpha-1}}.$$

If this condition holds, which the proposition assumes, then $\frac{d\Delta Y}{d\lambda} = \frac{d\Delta Y_p}{d\lambda} + \frac{d\Delta Y_c}{d\lambda} < 0$ for all λ. □

A.6 Calibration of the Model

I compute the equilibria of the model under autarky and trade with the preliminary calibration displayed in Table A1.

Table A1: Calibration for Figures

Parameter	Description	Source	Value
φ_{min}	Pareto location parameter	(Melitz and Redding, 2015, p. 1128)	1
θ	Pareto shape parameter	(Melitz and Redding, 2015, p. 1128)	4.25
α	Curvature in production	(Gáfaro and Pellegrina, 2022, p. 14)	0.7
ℓ_c	Collective sector farm size	Census Data (1991, see notes)	0.0056
p_A	Export premium	(Gáfaro and Pellegrina, 2022, p. 15)	1.32
f_A	Cash crop fixed cost	(Gáfaro and Pellegrina, 2022, p. 15)	2.06
L	Supply of land	Arbitrary (equivalent to choosing units)	1

Notes: In the typical enumeration area, which roughly corresponds to a local land market, the mean farm size in the collective sector is 6 hectares, and the total supply of land is 1,079 hectares. In the model the latter is standardized to $L = 1$, and dividing the mean collective farm size by total land supply yields $\ell_c = 6/1079 \approx 0.0056$.

B Appendix to Data Section

B.1 Details on the Algorithm for Irrigation Pond Detection

The pond detection algorithm consists of the following steps:

1. **Labeling and Splitting.** I create a "ground truth" 2024 dataset by manually labeling all ponds in one of Michoacan's 113 municipalities (Ario), using very high-resolution 2024 panchromatic imagery from Google Earth (0.4m). I create an additional set of labels, for 1996, using 2m black-and-white aerial imagery from INEGI. The resulting dataset of labeled polygons for Ario, together with all the satellite imagery was then partitioned into disjoint training (80%) and validation (20%) sets.

2. **Training the Modern Detection Model.** A U-Net segmentation model was trained on 10m resolution, multispectral Sentinel-2 imagery from 2024. The model was trained exclusively on the 80% training split of the Ario data to predict the presence of ponds at the pixel level. The combination of spectral bands used as input included the green band of visible light, near-infrared (NIR), and shortwave-infrared bands (SWIR2), which are informative for water detection.

3. **Large-Scale Inference for 2024.** The trained U-Net model was deployed across the entire state of Michoacan to generate pixel-level probability maps for all available Sentinel-2 imagery (currently 2019, 2021, 2024). This provides a comprehensive, state-wide map of pond likelihood for the modern era.

4. **Thresholding and Polygonization.** An optimal probability threshold was determined by maximizing the F1-score (the harmonic mean of precision and recall) on the held-out 20% validation set from Ario. This threshold was then applied to the state-wide probability maps to create a binary prediction of pond presence for each pixel.

5. **Vectorization.** Sets of contiguous positive pixels from the binary maps were converted into vector polygons. Each polygon represents a single, detected pond, forming the final dataset of modern agricultural water reservoirs.

6. **Training the Historical Classifier.** To determine the historical presence of the 2024-detected ponds, a separate classification model was trained. This model, a lightweight U-Net, takes as input 500×500 meter scenes extracted from 1996 Landsat imagery. The scenes are centered on the locations of known 2024 ponds, and the model is trained on the manually-created 1996 Ario labels to classify whether a given scene contained a pond in that historical year. The trained model uses the green band of visible light , the two NIR bands provided by Landsat-5, MNDWI, and the normalized difference vegetation index (NDVI) as input features. These bands were chosen based on their effectiveness in distinguishing water bodies from other land cover types in historical imagery.

7. **Historical Inference (1985–2018).** Finally, the trained historical classifier is used to perform inference across the full historical period. For each pond polygon detected in 2024, a corresponding scene is extracted from the historical Landsat 5 (1985–2012) archives for each year. The classifier predicts the probability of pond presence, effectively propagating backwards the machine learning predictions of pond presence through time.

Table B1: Pond construction cost benchmarks

Source	Total Cost (USD)	Notes
Ag. Dept. of Hidalgo (source)	$314,234	Unspecified size and investment is "over 4MM" MXN
FAISMUN 2025 (source)	$161,244	Unspecified size
Finance Ministry (source)	$135,072	Assumed 40x28m size
National Arid Zones Commission (source)	$108,534	Unspecified size
FAISMUN 2025 (source)	$94,056	Unspecified size
FAISMUN 2025 (source)	$84,710	Unspecified size
Ag. Ministry (SAGARPA) (source)	$80,892	Assumed 40x28m size
Finance Ministry (source)	$79,440	Assumed 40x28m size
Ag. Development Ministry (SEDAPA) (source)	$71,311	20x20x2m
Municipality of San Francisco Teopan (Oaxaca) (source)	$70,413	Unspecified size
Social Development Ministry - Municipality of Vanegas (source)	$68,885	Unspecified size
Ag. Development Ministry (SEDAPA) (source)	$64,700	Unspecified size
National Arid Zones Commission (source)	$61,532	Unspecified size
State Govt. of Oaxaca (source)	$57,774	Unspecified size
Official Gazette of Michoacan (source)	$41,538	Assumed 40x28m size
FAISMUN 2025 (source)	$39,960	Unspecified size
MACMO (direct contact)	$35,787	Unspecified size
CMCG Construction (source)	$23,936	2,800m^3 pond, 4m depth
Public Works, Municipality of San Juan del Rio (source)	$23,159	Unspecified size
Ag. Development Ministry (SIA) (source)	$20,018	Farmer subsidy cap 80%
FAISMUN 2025 (source)	$18,036	Unspecified size

Continued on next page

(continued)

Source	Total Cost (USD)	Notes
FAISMUN (Campeche) (source)	$16,746	Unspecified size
Public Works, Municipality of San Juan del Rio (source)	$16,170	Unspecified size
FAISMUN (Campeche) (source)	$14,886	Unspecified size
Finance Ministry (Michoacan) (source)	$13,986	Unspecified size
FAISMUN (Campeche) (source)	$13,025	Unspecified size
FAISMUN (Campeche) (source)	$11,164	Unspecified size
CMCG Construction (source)	$9,574	2,800m^3 pond, 4m depth
Mean	$62,528	

Notes: All costs include full pond construction (excavation, geomembrane, installation). N=28 observations.

C Additional Figures

Figure C1: Receiver Operating Characteristic (ROC) curves for pond detection models

(a) Modern pond detection model (Sentinel-2) **(b)** Historical pond detection model (Landsat 5)

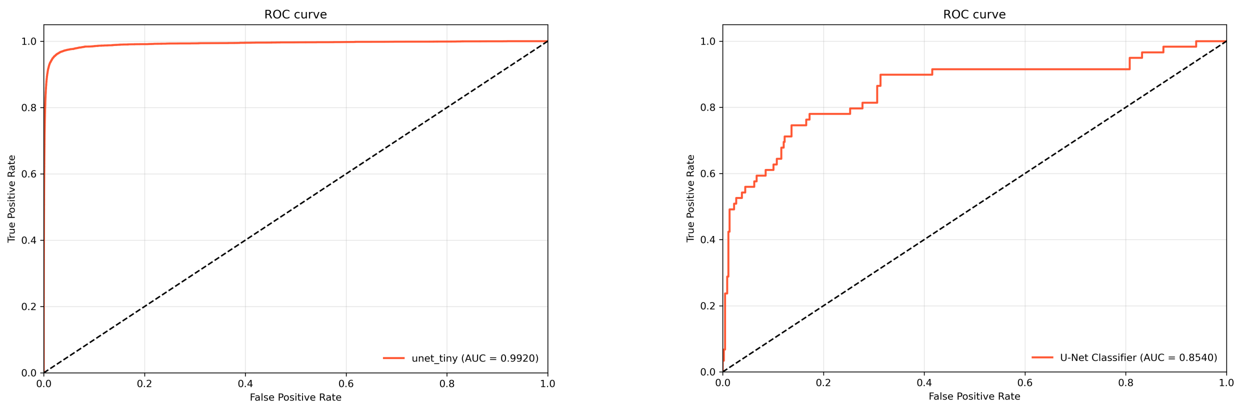


Figure C2: Census microdata descriptives, Michoacan

(a) Value per hectare by farm size (2007) **(b)** Avocado share by farm size (2007)

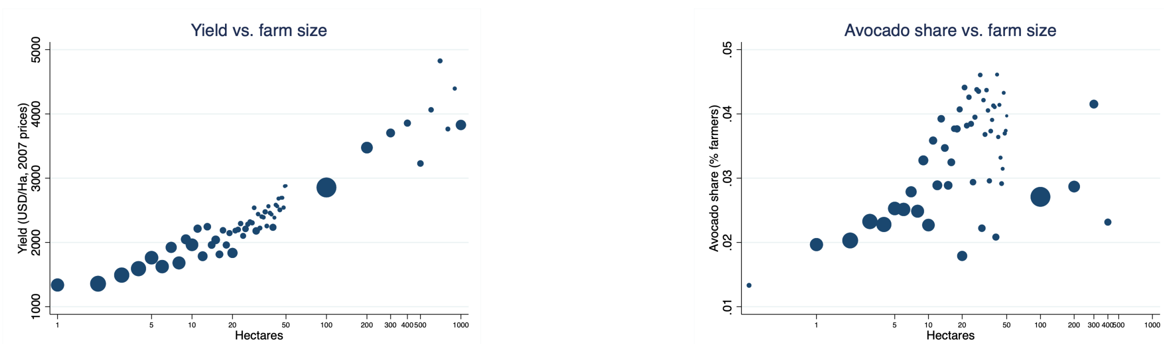
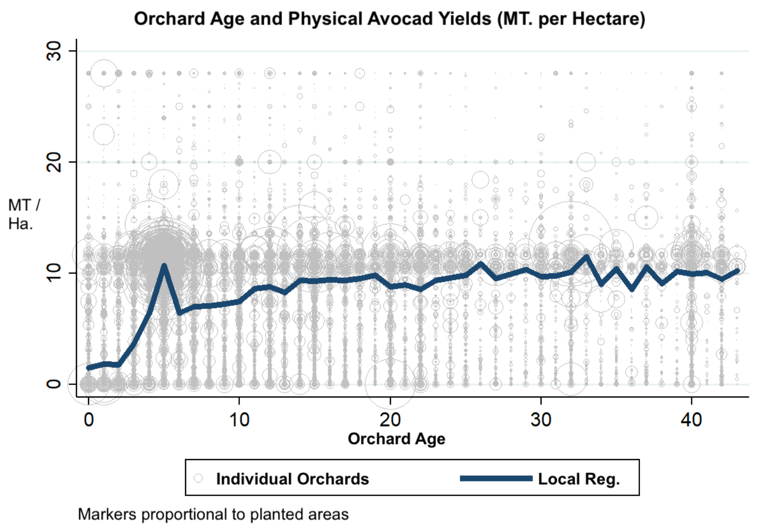
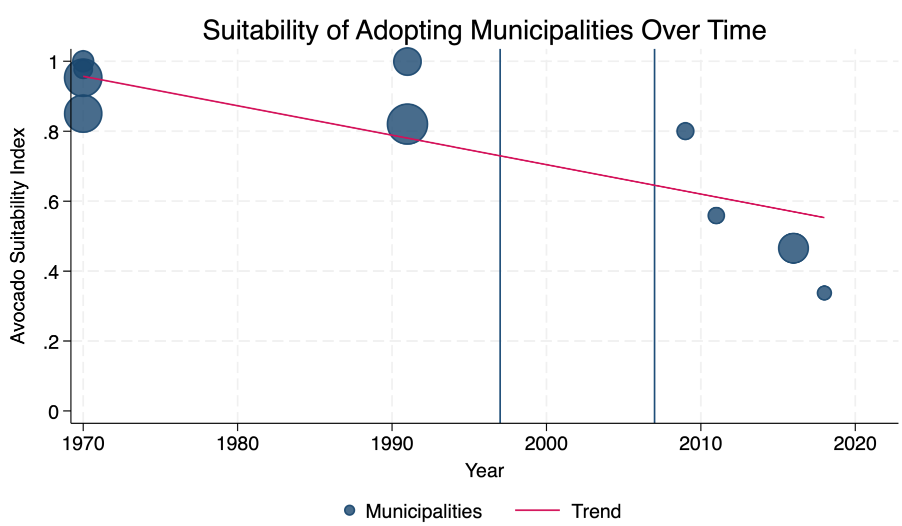


Figure C3: Avocado yields in Mexico (tons per hectare) as a function of orchard age



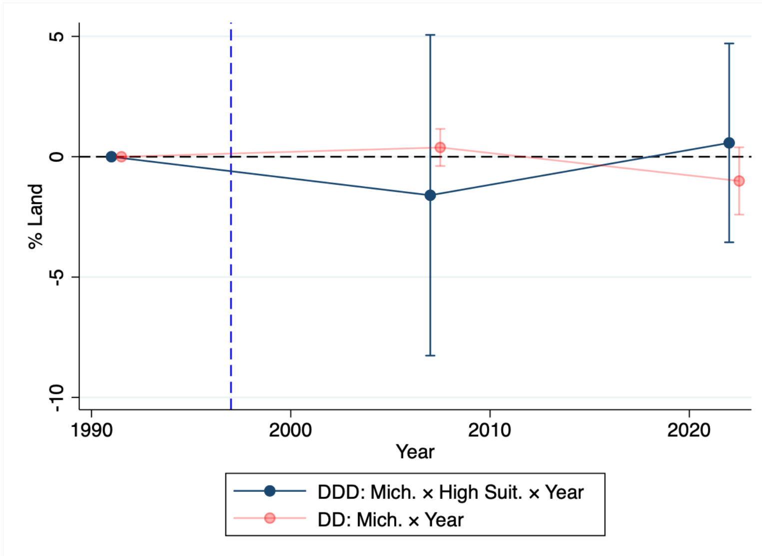
Notes: The figure displays average avocado yields, in tons per hectare, as a function of orchard age. The data comes from the 2022 Agricultural Census and covers the universe of avocado orchards. The dots, proportional to orchard size in planted hectares, represent individual orchards. The bold line presents local regressions estimated with Stata's `npgregress` command and default settings.

Figure C4: Declining suitability of municipalities switching to avocado cultivation over time



Notes: The figure displays the average suitability index of municipalities that "switch" to avocado cultivation in each year from 2000 to 2018. I consider a municipality switched in the first year in which its land area planted with avocados exceeds 5% of its total agricultural land area. The suitability index is constructed based on soil type, slope, temperature, and precipitation, following De Haro (2022).

Figure C5: Placebo test: land share planted by non-avocado farms larger than 100 hectares



Notes: The figure displays estimates of equation 10 using census data from 1991 (baseline), 2007, and 2022 at the municipality level. The dependent variable is the share of land planted, in each municipality, by farms above 100 hectares that do not grow avocados. Error bars represent 95% confidence intervals clustered at the municipality level.

D Additional Tables

Table D1: Descriptive Statistics by Treatment Status (EA Level)

Treatment Status	All	Treated	Control	Control	Control
Treated State	-	Yes	Yes	No	No
Avocado Suitability	-	Yes	No	Yes	No
Panel A: Land Use					
Total Hectares	1,766.07	1,270.56	1,621.14	1,546.36	2,007.02
	(1,809.92)	(1,190.50)	(1,419.16)	(1,803.86)	(1,898.84)
Mean Farm Size (Ha.)	11.72	9.50	19.71	10.59	11.81
	(20.20)	(12.58)	(28.76)	(21.39)	(18.73)
Ejido Share (%)	43.05	33.24	40.33	49.84	40.36
	(34.53)	(31.59)	(32.44)	(36.43)	(33.13)
Panel B: Land Consolidation					
Share Land Above 100 Ha. (%)	2.76	1.79	5.76	2.37	2.78
	(7.23)	(5.15)	(12.91)	(6.64)	(6.77)
Share Land Below 5 Ha. (%)	69.79	68.87	62.76	74.98	67.36
	(25.40)	(22.46)	(26.97)	(25.23)	(25.15)
Panel C: Suitability					
Avocado Suitability	0.35	0.72	0.22	0.65	0.11
	(0.31)	(0.22)	(0.09)	(0.19)	(0.14)
Panel D: Avocado Ponds					
Pond Coverage (%)	0.22	0.12	0.34	—	—
	(0.55)	(0.21)	(0.76)	—	—
Enumeration Areas	2,419	190	157	830	1,242
Observations	7,257	570	471	2,490	3,726

Notes: Panel A variables: Total Hectares - Total agricultural land (Ha.); Mean Farm Size - Average farm size in hectares; Ejido Share - Share of land in collective ejido tenure (%). Panel B variables: Share Land Above 100 Ha. - Share of privately owned land in parcels above 100 hectares (%); Share Land Below 5 Ha. - Share of privately owned land in parcels below 5 hectares (%). Panel C variables: Avocado Suitability - Continuous measure of avocado suitability. Panel D variables: Pond Coverage - Percentage of agricultural area covered by avocado ponds (%). Treated state refers to Michoacan. Avocado suitability is defined as above median suitability (weighted by 1991 hectares) within Michoacan. Standard deviations in parentheses. Unit of observation: Enumeration Area (EA). Data years: 1991, 2007, 2022 (agricultural census).