



Department of Economics
Working Paper Series

2020/002R

Crop Choice, Trade Costs, and
Agricultural Productivity

Alberto Rivera-Padilla

January 2020; Revised May 2020

Crop Choice, Trade Costs, and Agricultural Productivity*

Alberto Rivera-Padilla[†]

May 2020

Abstract

I argue that the agricultural productivity puzzle is in large part a staple productivity puzzle. Using detailed data from Mexican farms, I show that most farmers grow staple crops, despite the fact that labor productivity in cash crops is substantially higher. To explain this pattern I develop a quantitative general equilibrium framework with multiple regions and crop types, subsistence requirements of staple food, and interregional trade costs. In equilibrium, most farming production is in staple crops because subsistence constraints and high trade costs prevent most farmers from specializing in cash crops. Reducing trade costs in Mexico to the U.S. level would raise the ratio of employment in cash crops to staples by 15 percent and generate a 13 percent increase in agricultural labor productivity.

JEL Codes: E01, E24, J43, J61, O11, O13, O18, O41, R11, R40

Keywords: agriculture, productivity, crop choice, trade costs

*I am grateful to Todd Schoellman for his guidance and support, and to Berthold Herrendorf, Domenico Ferraro and Gustavo Ventura for their valuable comments and suggestions. I thank workshop participants at Arizona State and the Econometric Society NASM for helpful comments and discussions, especially Richard Rogerson, Bart Hobijn, and Alexander Bick. I thank the editor and two anonymous referees for their insightful comments and suggestions. Lastly, I owe special thanks to Natalia Volkow and the personnel from the Microdata Laboratory of INEGI. The usual disclaimer applies.

[†]California State University Fullerton, Fullerton, CA 92834-6848. Email: ariverapadilla@fullerton.edu

1 Introduction

A large body of literature documents that studying agriculture is critical for understanding cross-country income differences. The reason is twofold. First, while poor countries are much less productive in aggregate output per worker, the productivity gaps are particularly large in agriculture. Second, despite these large productivity gaps, poor countries allocate a high share of their labor force to agriculture (Caselli, 2005; Restuccia et al., 2008). Combined, these two facts prompt a key question: why do poor countries devote so much labor to such an unproductive sector?

The main goal of this paper is to further refine this puzzle. I use detailed, restricted access data from Mexican farms to document two facts based on comparing staple crops, such as maize, and fruits that are usually grown as cash crops. First, while Mexican agriculture is much less productive than non-agriculture, the productivity gaps in value added per worker are much larger for staple crops. Second, despite these large productivity differences, farm labor in Mexico is mostly devoted to staples production. Together, these two facts suggest that we can focus on an even narrower question: why do poor countries devote so much labor to unproductive staple crops? If Mexico had the agricultural employment shares in the United States for grains and fruits, fixing productivity in each of these sectors, value added per worker in agriculture would be 14 percent higher.

This paper proposes an explanation based on two key mechanisms that determine the efficient crop choice by farmers: subsistence requirements of staple crops and *interregional* trade costs. The former is based on the fact that staple crops have a high caloric content and represent a crucial component of the population's diet. Moreover, staple crops are an important nutritional source for poor farmers who have incentives to produce their own food. This is supported by evidence from Mexican farms which shows that the share of production used for family consumption is much larger for staples. The second mechanism is based on the fact that farm-to-market trade costs are higher for fruits than for staple crops. This implies that fruit farmers receive a smaller share of the market value and must offer relatively low

prices to be competitive.

To formalize the analysis, I build a general equilibrium model with interregional trade and self-selection of farmers into types of crops. The model features non-homothetic preferences, costly trade across regions, and the existence of two agricultural goods: a staple crop (maize) and a cash crop (fruit). In this framework, workers move between urban and rural regions, and farmers choose to produce either type of crop as an efficient response to subsistence requirements of staple food and trade costs. The model highlights that productivity gains in the economy could lead to reallocation of employment *within* agriculture. This is different from most of the literature on structural transformation which focuses on the relationship between sectoral productivity gains and reallocation of labor from agriculture to non-agriculture (see e.g. [Herrendorf et al., 2014](#)).

I calibrate the model to match features of the Mexican economy. For the topic of interest in this paper, Mexico is poor enough and the share of labor in agriculture (14 percent in 2014) is large compared to rich countries (1 percent in the United States). Furthermore, I make use of rich disaggregate data from Mexico that are not typically available in low-income countries. In particular, I use detailed farm data on prices, production value, expenses, employment, and land usage at the crop level. I estimate trade costs from price gaps of homogeneous goods across regions in the country; thus, the definition of trade costs is broad and includes more than just transportation costs between distant regions, they also represent possible monopoly power of intermediaries. The quantitative results of the model imply that trade costs can account for a considerable proportion of the relative employment between maize and fruits, and a lower labor productivity in agriculture. In a counterfactual case without trade costs, agricultural labor productivity increases by 21 percent and the ratio of employment in fruits to maize increases by 17 percent.

A key quantitative feature of the model is that it generates differences in value added per worker and value added per hectare across types of crops without introducing a wedge in marginal products or implying implausible differences in factor income shares. The calibration implies that productivity differences between maize and fruits are due to the fact that most

farmers with low-productivity land self-select into staple crops. Indeed, while fruit producers tend to have highly productive land, a large fraction of maize farmers have a relatively low productivity in that sector.

This paper is related to the macroeconomic literature that has tried to explain agricultural productivity differences across countries (see e.g. [Gollin et al., 2014b](#)).¹ Recent papers in this literature have taken into account production decisions within agriculture, for instance, [Adamopoulos and Restuccia \(2020\)](#) develop a model that features crop choice to analyze the effects of land reforms on farm size and agricultural productivity. The model in this paper builds on the framework of [Lagakos and Waugh \(2013\)](#) and introduces a new margin to think about productivity gains in agriculture, namely, the choice between producing staple or cash crops. Their selection model augmented with different agricultural goods and interregional trade allows me to quantify the importance of trade costs for labor allocations across types of crops.

The puzzling concentration of Mexican agricultural labor in low-productivity staple crops is connected with the fact that subsistence farming is mostly labor intensive; in contrast, production of the same staple grains is highly mechanized in richer countries such as the United States. Therefore, this work relates to recent papers showing that differences in farming mechanization and capital per worker are important to explain cross-country differences in agricultural labor productivity ([Caunedo and Keller, 2019](#); [Chen, 2020](#)).

Additionally, this paper is closely related to recent literature that examines the effects of transportation costs on interregional trade and welfare using general equilibrium models (see e.g. [Donaldson and Hornbeck, 2016](#); [Donaldson, 2018](#); [Alder, 2019](#)). I build on the methodology of this literature to measure sector-specific trade costs across regions, and construct a unique dataset that combines farm and market data to compare prices of homogeneous goods between origins and destinations in Mexico.

¹Various explanations include inefficient factor markets ([Vollrath, 2009](#)), distortions related to farm size ([Adamopoulos and Restuccia, 2014](#)), risk and low use of intermediate inputs ([Donovan, 2018](#)), and misallocation of land ([Chen, 2017](#); [Adamopoulos and Restuccia, 2018](#); [Gottlieb and Grobovšek, 2019](#); [Le, 2020](#)). A related literature focuses on the productivity gap between agriculture and non-agriculture within countries (e.g. [Gollin et al., 2014a](#); [Herrendorf and Schoellman, 2015](#)).

Lastly, this is not the first paper that studies the relationship between trade, agriculture, and development ([Adamopoulos, 2011](#); [Herrendorf et al., 2012](#); [Tombe, 2015](#)). In particular, [Gollin and Rogerson \(2014\)](#) build an interregional trade model in which subsistence agriculture is associated with remote rural areas, and [Sotelo \(2019\)](#) studies the effects of regional trade frictions on welfare and farm productivity in Peru. A key difference with these papers is that I study gaps in labor productivity between two distinct categories of crops, staples and fruits, to show that low agricultural productivity is largely driven by a high allocation of resources to low-productivity staple crops. In the model, subsistence requirements of staple food imply that trade costs amplify the selection of farmers into staple crops by reducing the relative price of fruits in rural regions. Thus, this paper relates subsistence agriculture to production of staples instead of high-value cash crops.

The rest of the paper is organized as follows. Section 2 presents a description of the microdata used in this paper, as well as empirical evidence on crop productivity and trade costs in agricultural markets in Mexico. Section 3 introduces a multi-sector selection model with interregional trade. Section 4 provides a description of the calibration and presents the quantitative results of counterfactual experiments. Section 5 provides a discussion of important assumptions in model and their possible implications. Finally, Section 6 concludes.

2 Empirical Evidence

To present the empirical evidence, I simplify the analysis by narrowing down the number of crops considered. Based on its production volume, harvested land, and relevance for subsistence, maize is the most important crop in Mexico. In this paper I will use it as a benchmark of staple crops and compare it to other fruits that are among the most important cash crops in the country.

I use restricted access, farm-level data from agricultural surveys in Mexico. These microdata is part of the *Encuesta Nacional Agropecuaria* (ENA) 2014. The surveys were taken from a sample of 75,148 farms in 25,800 localities of the country during the agricultural cycle from

fall 2013 to fall 2014. They gathered nationally representative data for 34 products that were chosen based on their contribution to Gross Domestic Product (GDP). The unit of observation in the survey is a unit of agricultural production formed by a set of plots located in the same municipality. Then, since more than one crop can be grown by a unit of production throughout the agricultural year, each observation in the database represents a farm-crop pair.²

The target population of the surveys were all the production units that reported data for one of the products of interest in the agricultural census 2007. According to the ENA 2014 surveys, 69 percent of irrigation farms (18 percent of all farms) and 71 percent of rainfall farms had at most 5 hectares (12 acres). Commercial farmers are defined as those with more than 5 hectares in irrigation farms and more than 20 hectares in rainfall farms. These farmers represent 10 percent of all farms and they own 57 percent of the land. This is consistent with the farm size distribution from the 2007 agricultural census: 66 percent of farms had no more than 5 hectares and almost 90 percent of farms had less than 20 hectares.

The ENA 2014 surveys have detailed information by variety of crop at the farm level. For each crop that is grown in a farm, the surveys report harvested land and production volume; amounts of production used for family consumption, feed, and seed; farm-gate prices of output sold; quantities used of fertilizers; and farmers' expenses in different stages of production. The latter include expenditures on modern inputs such as chemicals, pesticides, and irrigation. In addition, hired labor and other farm expenses are reported at the farm level. The fact that most expenditures are reported at the crop level allows me to calculate valued added for each of them. The latter is key to make productivity comparisons across crops.

In addition to the farm data, I use national accounts data and various sources of agricultural data provided by Mexico's government agencies. Specifically, I use data from the *Sistema Nacional de Información e Integración de Mercados* (SNIIM) to get data on wholesale prices for particular varieties of crops in every state of the country. See Appendix A for more details.

²These data is subject to confidentiality regulations and access is granted on-site at government offices located in Mexico City. These data is provided by the *Sistema Nacional de Información Estadística y Geográfica* (SNIEG) of the *Instituto Nacional de Estadística y Geografía* (INEGI). The views and conclusions expressed are exclusive of the author and do not reflect official positions or statistics of SNIEG, or INEGI.

2.1 Crop Productivity

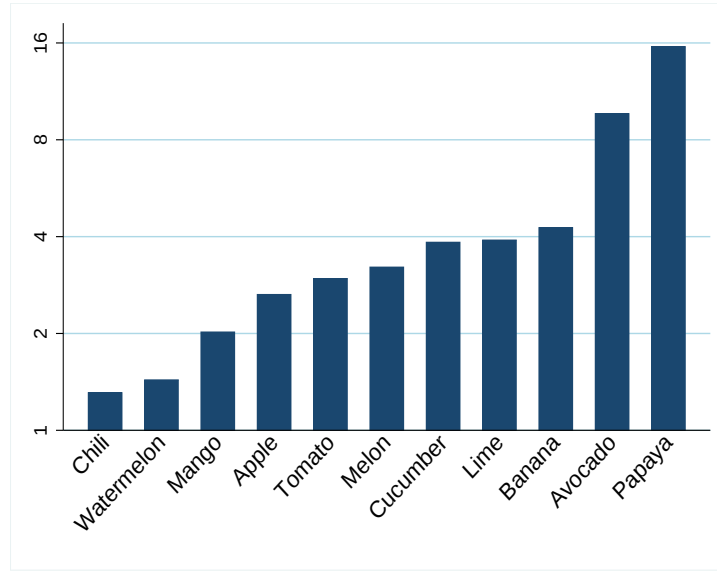
In this section I show that productivity in maize farming is significantly lower than productivity in fruit farming. I consider two measures of crop productivity: value added per worker and value added per hectare. I calculate value added for each farm-crop observation in the following way. First, I obtain the value of production net of the amount used for seed and animal feed; then, I subtract expenditures on fertilizers, pesticides, and irrigation. The amount of labor is reported at the farm level, thus, I focus on farms producing one type of crop to measure labor productivity. The results from this section are complemented in Appendix C using alternative sources of public aggregate data.

Panel A in Figure 1 shows aggregate value added per worker for different fruits relative to value added per worker in maize farming. For most of these fruits labor productivity is over two times higher than maize, and the average productivity gap is around six. In comparison, the ratio of value added per worker in non-agriculture to agriculture was 5.7 in Mexico in 2013. Thus, the agricultural productivity gap in Mexico has a similar magnitude to the productivity gap between these fruits and maize. However, despite these productivity differences, Panel B shows that labor allocated to maize is much higher than any of the fruits considered; in fact, all these crops together add up to 46 percent of total workers allocated to maize.

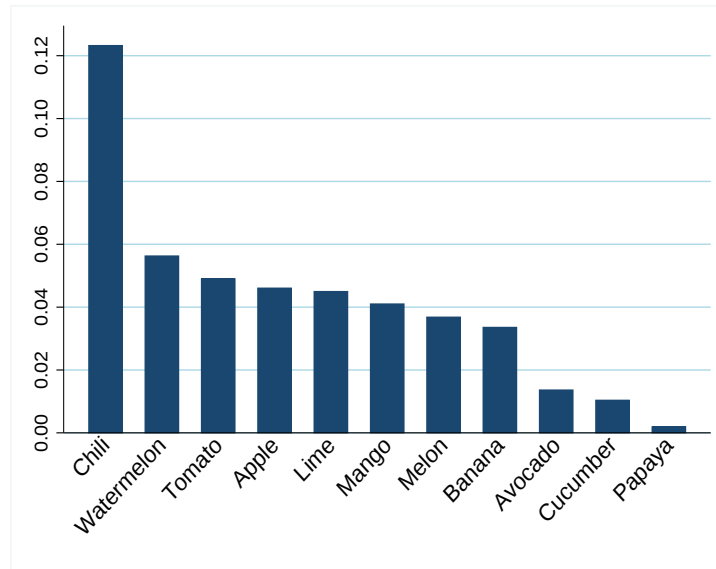
Moreover, given that labor productivity in agriculture is equal to the weighted sum of labor productivity in each sub-sector, a large employment share in highly unproductive crops decreases the value of agricultural labor productivity in the economy. Then, these facts imply that the puzzling relationship between high employment and low agricultural productivity in developing countries is in large part a puzzle of high employment in low-productivity staple crops. In other words, labor productivity in most of those fruits is so large compared to maize that it is puzzling such a small fraction of labor is allocated to them.

To put this evidence from Mexico into perspective, I compare it with the United States using data from the 2017 U.S. census of agriculture. The United States is not only a benchmark for rich countries, but the comparison with Mexico is convenient since both countries follow

Figure 1: Panel A. Value added per worker relative to maize



Panel B. Total workers relative to Maize



Notes: Maize is normalized to one in both cases. Total workers include family members participating in farming activities.
Source: Author's estimates using data from SNIEG and INEGI: ENA 2014.

the North American Industry Classification System (NAICS).³ Thus, I use national accounts data to compare total employment (hired and unpaid workers plus farming operators) in broad categories of crops.

Table 1 presents the share of workers out of total employment in agriculture for two cat-

³To the best of my knowledge, there is no public database with data on employment at the crop level across countries.

Table 1: Comparison of United States and Mexico

Country	Share of agricultural employment	
	Grains	Fruits & Vegetables
Mexico	35	19
United States	21	25

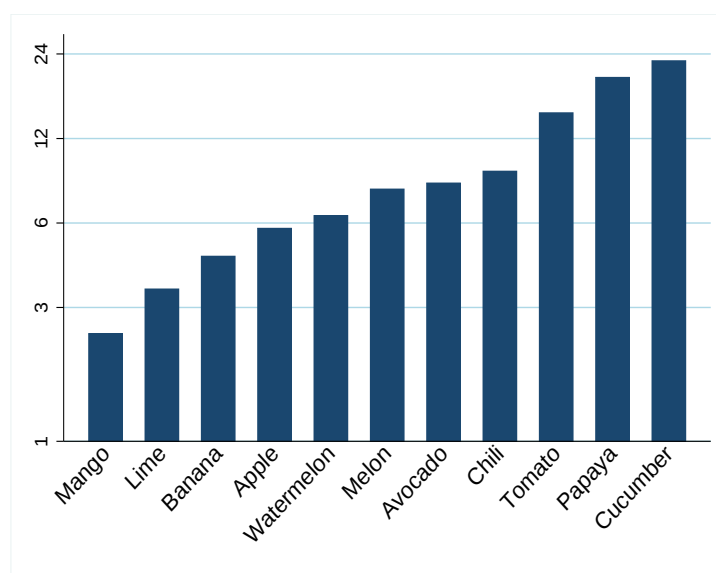
Notes: This table presents employment shares (percent) within agricultural crop production considering hired and unpaid workers, and farming operators. Grains corresponds to NAICs classification *Oilseed and grain farming (1111)*. Fruits and Vegetables includes *Vegetable and melon farming (11121)* and *Fruit and tree nut farming (1113)*.

Source: Author's estimates using data from Mexican 2013 Input-Output Tables and 2017 U.S. Census of Agriculture.

egories: (i) grains farming and (ii) fruits and vegetables farming. These numbers show that the share of agricultural employment in grains is higher in Mexico than in the United States, whereas the reverse is true for fruits and vegetables. The United States is the largest producer of maize in the world and, compared to Mexico, maize farming is highly mechanized. Thus, it suggests that as countries get richer and agriculture transitions from traditional to modern techniques, fruits farming increases its labor force relative to grains farming given that the former requires labor activities that are difficult to replace with machines such as picking and stacking.

I now use these data to do productivity accounting in the following way. I take the shares of agricultural employment in grains and fruits farming from the United States and use them to weight agricultural value added per worker in Mexico. I focus on crop production and adjust the employment shares proportionally so that the share in other types of crops remains the same. To be clear, I only adjust employment in grains and fruits according to their proportion in U.S. agriculture and fix value added per worker in each sub-sector. The results of this counterfactual exercise imply a 11 percent increase in agricultural labor productivity in Mexico, but if other types of crops are omitted the boost is equal to 14 percent. Therefore, if Mexican agricultural labor was reallocated from grains to fruits to match their proportion in U.S. agriculture, holding all else constant, there would be significant gains in value added per worker in agriculture.

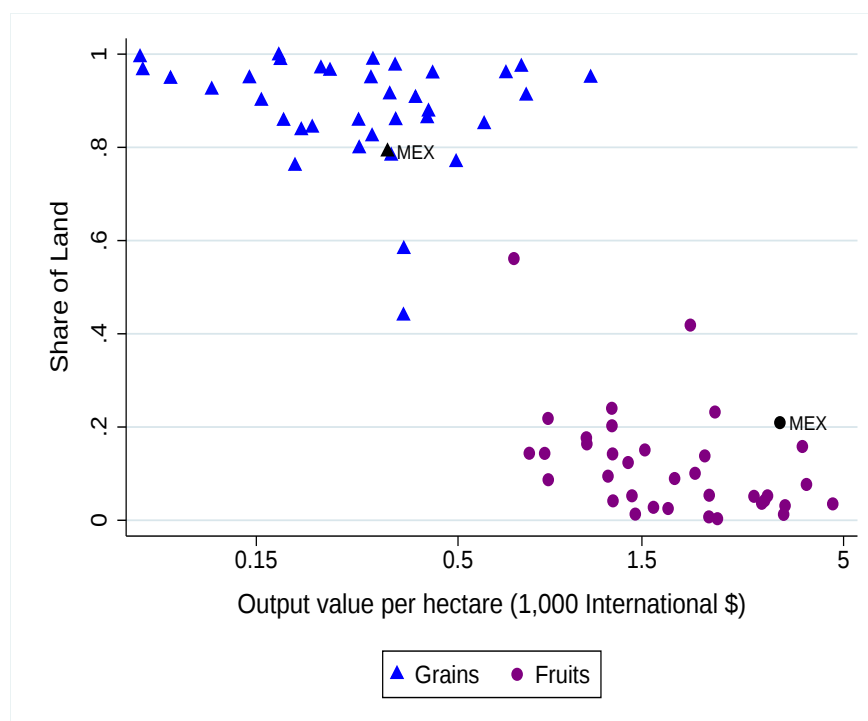
Figure 2: Value added per hectare relative to maize



Notes: Maize is normalized to one.

Source: Author's estimates using data from SNIEG and INEGI: ENA 2014.

Figure 3: Land allocation and yields in poor countries and Mexico



Notes: Average yields are reported for each category in a country. Yields are measured as net production value per hectare and are weighted by harvested hectares. The share of land is reported with respect to total harvested hectares in both categories. Output value (in constant 2004-2006 1,000 I\$) is defined as gross output value net of agricultural inputs (seed and feed). Includes Mexico and 35 countries classified as least developed by the United Nations.

Source: FAO, 2010.

Next, Figure 2 presents the gaps in aggregate value added per hectare for the same group of crops. The differences in land productivity between maize and fruits are even larger than the labor productivity gaps. In this case, I can use data from the Food and Agriculture Organization (FAO) to compare Mexico with poorer countries. Figure 3 compares yields and harvested land between grains and fruits using such data.⁴ Almost every poor country allocates a large share of their land to produce staple crops, even though yields of fruits are significantly higher. While similar patterns can be found in richer countries, there are two reasons these facts are especially important in poor countries: (i) a high share of the population works in agriculture, which means a large share of the labor force produces low-productivity crops; and (ii) producing fruits involves labor intensive activities, thus, poor countries could exploit their production given the capital constraints, small farm size, and labor intensive techniques of most farmers.

One possible concern with the aggregate results described in previous paragraphs is that such productivity gaps are driven by differences in farm size between maize and fruit producers in Mexico, or by particular regions of the country that are highly productive in fruit farming. To address these issues, I estimate productivity gaps between types of crops controlling for state and farm size. The results presented in Table 2 show that productivity is significantly larger for fruits than for maize even if such controls are taken into account. That is, adjusting for region and the size of farms, the labor productivity gap between fruits and maize is 3.4 (the raw gap is 5.8), while the land productivity gap is 4.5 (the raw gap is 4.8). Therefore, the adjusted productivity gaps are smaller, but still sizable.

The empirical results presented in this section suggest the idea that a significant fraction of farmers must be relatively unproductive at producing staple crops: not every farmer has the best land to grow maize, nor the set of skills or knowledge required to produce such type of crop. The fact that most farmers decide to grow staple crops implies that there might be

⁴This figure shows output value per hectare. In poor countries, this could be a good approximation to value added per hectare since intermediate inputs usage is low; however, comparing land productivity across crops, especially in rich countries, would require value added at the crop level. To the best of my knowledge, there is no public source with such data.

Table 2: Productivity gaps: fruits relative to maize

Value Added per Worker	Raw Gap	Adjusted Gap
Gap	5.8	3.4
State fixed effects, farm size	No	Yes
Observations	26,197	26,197
Value Added per Hectare	Raw Gap	Adjusted Gap
Gap	4.8	4.5
State fixed effects, farm size	No	Yes
Observations	33,189	33,189

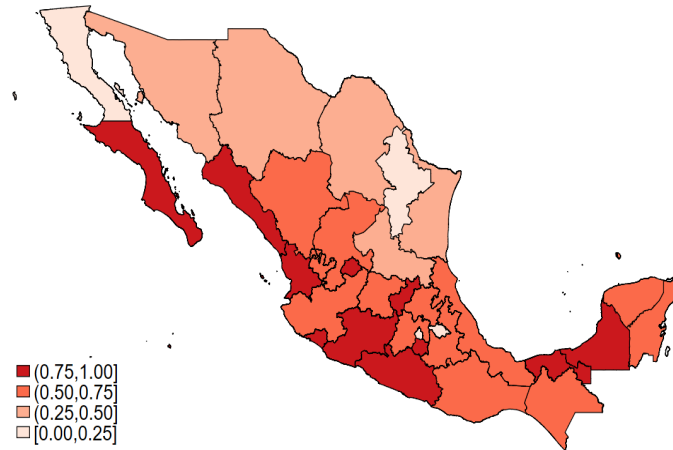
Notes: Results obtained from regressing $\log(\text{value added per worker})$ and $\log(\text{value added per hectare})$ on a dummy that takes a value of 1 for fruits and 0 for maize. The gap reported is the exponential of the estimated dummy coefficient. Controls include (log) agricultural land of the farm and state dummies. The coefficient of fruits is significant at the 1 percent level in every case. Regressions are weighted by worker and hectares, respectively.
Source: Author's estimates using data from INEGI-ENA 2014.

barriers amplifying the selection of farmers into those crops. That said, one might naturally wonder if maize farmers can actually grow a more productive crop in their location. To address this concern, I use aggregate public data at the municipality level to take a comprehensive look at actual yields value of fruits and maize across the country. The goal is to show that in most Mexican regions there is production of high-productivity cash crops, so maize is not the only feasible option for farmers.

Figure 4 presents a map with the fraction of municipalities in each state that have a higher average output value per hectare in fruits than maize. The median share of municipalities across states is equal to 69 percent and the remaining fraction are mostly cases where either maize or fruits did not have production.⁵ Thus, while specific subregions of the country might attain higher fruit yields due to better geographical suitability, in most regions farmers seem to have the possibility to produce some high-productivity fruits. Indeed, it is not the case that only one region of the country grows fruits with higher yields than maize, which indicates that

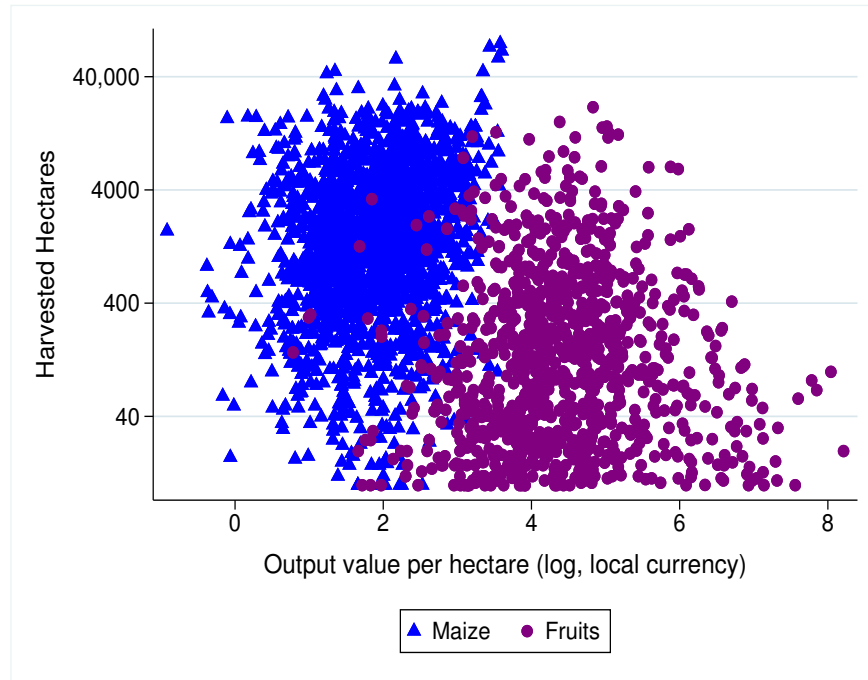
⁵States in the north of the country have the lowest share. A possible explanation is that a large portion of the north region is arid and agro-climatic conditions are less favorable, so production is concentrated in particular locations.

Figure 4: Share of municipalities with high-productivity fruits



Notes: This figure shows the fraction of municipalities in each state where average output value per hectare is larger for fruits than for maize. There are 2,457 municipalities and 32 states in the country.
Source: Author's estimates using SIAP data, 2014.

Figure 5: Land allocation and yields across municipalities in Mexico



Notes: This figure presents output value per hectare and harvested land at the municipality level.
Source: Author's estimates using SIAP data, 2014.

in many areas farmers are actually choosing to grow maize even if this crop is not the only feasible option. In other words, this evidence suggests that farmers' inability to specialize in more productive crops is the most plausible story, rather than inability to grow them. See Appendix B for additional evidence in particularly poor states.

Lastly, using the same data from the previous map, Figure 5 plots yield values and harvested land for maize and fruits across municipalities in Mexico. There are two key observations in this figure. First, it displays the large number of localities in the country where farmers can grow high-productivity fruits. Second, the figure shows that for every level of land yields of fruits are significantly higher, yet most of the land is allocated to maize.

2.2 Trade costs

This section presents evidence that trade costs are large in agricultural markets in Mexico. I measure these costs indirectly using differences in prices across regions. Thus, trade costs consist of more than just moving goods across distant regions, instead they reflect the quality of infrastructure for transportation and storage in each region, and how competitive markets are. My data satisfies two important characteristics to measure trade costs using spatial price gaps: (i) homogeneous products, and (ii) regions that are actually trading with each other. A similar empirical strategy is used by [Donaldson \(2018\)](#) based on regional varieties of salt in India.

First, to measure trade costs of crops, I compare farm-gate prices with wholesale market prices across states in 2014. I build a unique dataset of prices for specific crop varieties by combining farm data from ENA 2014 surveys with prices listed in SNIIM in the same year. These prices come from large supply centers of agricultural products located in the major cities of every state.⁶ The latter are reported monthly, so I calculate the average price for each variety of crop in every market by origin. Market prices only specify the state of origin for each product, so I also aggregate farm prices to the state-level. Price gaps between each origin and destination are measured as the farm price divided by the wholesale price (which I define as the farm share). I only consider origins and destinations that are potentially trading, that is, observations in which the destination had a price greater or equal than the origin price. After

⁶In most cases, the main market is located in the capital of the state. In states with multiple markets, I took the one that seemed to have more trade based on the information. Agricultural goods in Mexico are mostly traded by road transport, so I do not consider the states of Baja California and Baja California Sur due to their relatively remote location. The results are not sensitive to the inclusion of these places.

Table 3: Trade costs of fruits relative to white maize

Crop (Variety)	Raw Gap	Adjusted by Routes
Avocado (Hass)	1.5	1.4
Cucumber	2.4	2.0
Pepper (Poblano)	2.1	1.9
Tomato (Saladette)	1.5	1.4
Watermelon (Rayada)	1.9	2.0

Notes: Adjusted gaps take into account origin-destination fixed effects. Origins and destinations represent states in Mexico, and comparisons are made between prices reported by farmers and wholesale prices in major cities.

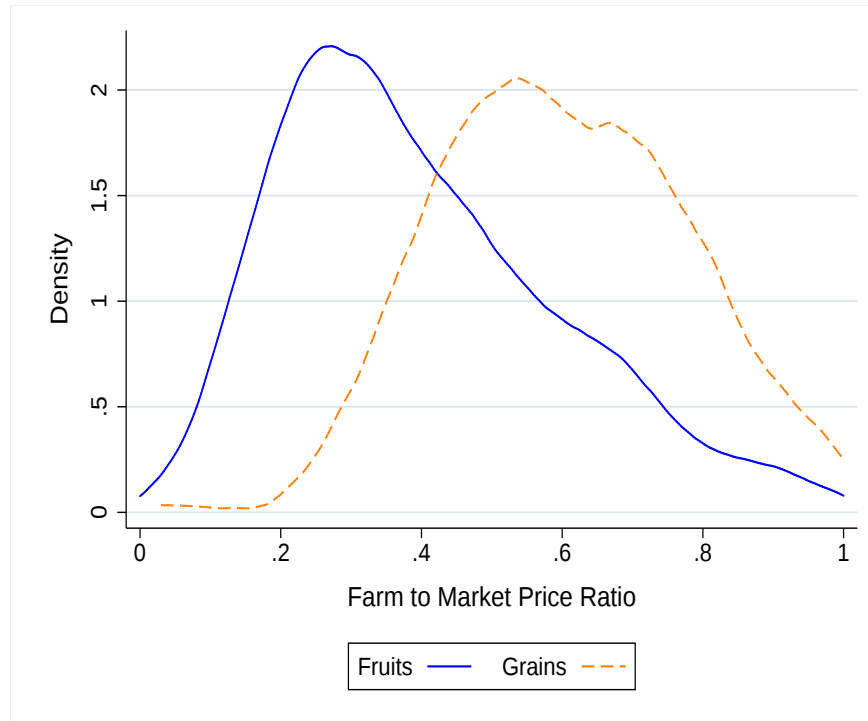
Source: Author's estimates using data from INEGI-ENA 2014 and SNIIM.

this, I end up with 930 origin-destination observations for 68 fruits and grains traded across 30 states.

Figure 6 presents the distribution of trade costs for grains and fruits. The median farm share is 59 percent for grains and 36 percent for fruits. This means that a typical grain farmer receives more than half of the wholesale value, while a fruit farmer receives a little more than a third of the value. Specifically, the median farm share of maize (*white*) is 69 percent; in comparison, the median farm share of avocado (*hass*) and pepper (*poblano*) are 49 and 32 percent, respectively. One possible concern with these results is that fruits might be traded to further distances than grains. In the second column of Table 3, I present the raw gap in trade costs between a selection of fruits and white maize, while the third column shows the estimated gap controlling for origin-destination fixed effects. There is almost no differences between raw and adjusted gaps; thus, the differences in trade costs are not driven by differences in trading routes.

These estimates are in line with previous studies on Mexico's agricultural sector according to which 30 percent of agricultural production is lost due to inadequate transport and storage facilities (see OECD, 2007). According to these studies, farmers in fruits and vegetables markets receive between 35 and 45 percent of retail prices; in comparison, producers in other Latin American countries receive 50 percent of retail prices, and in some cases of Central America between 65 and 75 percent. The existence of few intermediaries controlling the distribution of

Figure 6: Distribution of trade costs by type of crop



Notes: Trade costs are equal to the reciprocal of the farm-to-market price ratio. Kernel densities (Epanechnikov). Based on 930 origin-destination observations for 68 varieties of fruits and grains traded across 30 states in Mexico. Origins and destinations represent states, and comparisons are made between farm prices and wholesale prices in major cities. Source: Author's estimates using data from INEGI-ENA 2014 and SNIIM.

cash crops was related to the low prices that farmers face in Mexico. This is also supported in my data by the fact that farm shares are low even within the same state: 47 percent on average versus 43 percent across states. The latter could also be explained by differences in remoteness and infrastructure quality within states, and by a large fixed cost component in transportation (e.g. trucks and facilities with refrigeration).

The previous paragraphs focused on trade costs of output in agricultural markets. However, trade costs of modern inputs are also relevant. To measure trade costs of fertilizers, I use a similar approach though the available data is different. First, from the ENA 2014 surveys, I obtain quantities in tons of chemicals and natural fertilizers used for crop production. Total expenses on fertilizers are also reported for each crop. Then, I calculate implicit prices of fertilizers at the farm level dividing total expenditures by total quantity of fertilizers. A wide range of varieties of chemical fertilizers are reported by farmers, so I focus on the chemicals

Table 4: Trade costs of fertilizers
Farm fertilizer prices relative to market price in origin

Variety	Price gap
Urea	1.62
Ammonium Sulfate	1.53

Notes: Median gap at the national level. Most of Urea is imported so the origin refers to a port (Veracruz State). For Ammonium Sulfate, origin refers to the region where production plants are located (Queretaro State). See text for details.
Source: Author's estimates using data from INEGI-ENA 2014, and SNIIM.

that account for the majority of the observations: Urea and Ammonium Sulfate. The former is mostly imported from Eastern Europe, while the latter is primarily produced in the central region of the country.

I use data from SNIIM to get the commercial price of these fertilizers in their possible state of origin: Veracruz (a port) for Urea and Queretaro (a production plant) for Ammonium Sulfate. To have a higher number of observations, I estimate trade costs for these products by aggregating farm prices to the municipality level and dividing them by the commercial price at the state of origin. Table 4 presents the median farm to market ratio. The magnitude of these gaps is consistent with the fact that most farmers report that a main obstacle of production are the high costs to acquire modern inputs.

This section provided evidence on the existence of large trade costs in agricultural markets in Mexico. Intuitively, these trade costs might amplify the number of farmers producing maize relative to fruits for the following reasons. First, farmers producing fruits receive a lower share of their market value because trade costs are higher in this sector. Second, when staple food is costly to trade from farms to dense urban regions, more labor needs to be allocated to its production to guarantee that demand is satisfied. Finally, given that modern inputs are costly to acquire, more labor needs to be allocated to produce enough staple food for the population. The following section introduces a model to assess the quantitative importance of trade costs for crop choices and agricultural productivity in a context with subsistence requirements of staple food.

3 Model

I develop a static general equilibrium model with interregional trade. The model includes production of different agricultural goods and features heterogeneous productivity across farmers, non-homothetic preferences, and trade costs. The framework builds on the selection model of [Lagakos and Waugh \(2013\)](#), the trade literature based on [Eaton and Kortum \(2002\)](#), and the interregional trade model of [Herrendorf et al. \(2012\)](#).

3.1 Environment

There is an urban region denoted by u and a rural region denoted by r . Regions are indexed by $j \in \{u, r\}$. Each region is populated by a household of size N_j . Individuals can move freely between regions so N_j is endogenous. There are three sectors in the economy: a nonagricultural good (n), and two agricultural goods: one is maize (m), a staple crop which is used for subsistence requirements, and the other is a fruit (f) or cash crop. These goods are indexed by $s \in \{n, m, f\}$. I assume that the urban region only produces non-agricultural goods, whereas the rural region only produces agricultural goods. The non-agricultural good is used as an input in agricultural production (e.g. chemicals and fertilizers) and interregional trade is restricted by sector-specific trade costs. The model abstracts from other factors that could be considered important for agricultural production decisions such as international trade, land markets, and risk. [Section 5](#) discusses these simplifications and the possible quantitative implications of taking them into account.

The details of the decision process are presented below. Here, I describe the timing of choices in the model. First, individuals choose to live either in the urban or rural household. Then, households pool income to maximize utility of their members by choosing consumption and, in the rural household case, allocating members as farmers (farm operators) or farm-workers. Finally, farmers decide to produce either maize or fruits based on their individual productivity to produce each crop.

3.2 Preferences and endowments

In both regions, individual preferences are defined according to the utility function

$$U(c_{jm}, c_{jf}, c_{jn}) = \epsilon_m \log(c_{jm} - \bar{m}) + \epsilon_f \log(c_{jf}) + \epsilon_n \log(c_{jn}) \quad (1)$$

where $\sum_i \epsilon_i = 1$, $\bar{m} > 0$ is the subsistence requirement of maize consumption, and c_{js} is consumption per capita of good s in region j .

Each individual is endowed with one unit of time that is supplied inelastically to the labor market. Additionally, the rural household is endowed with L units of land to be used by farmers for agricultural production. Each farmer i is endowed with the same fraction of land ℓ and a pair of efficiency units of land $\{z_m^i, z_f^i\}$ to produce crops m and f , which is drawn from a distribution $G(z_m, z_f)$. However, a farmer can only produce one type of crop in her plot of land. The heterogeneity in land augmenting productivity across farmers can be interpreted as differences related to both the quality of land and the skills or knowledge of farmers to produce a crop.

3.3 Production technologies

The non-agricultural good is produced according to a constant returns to scale production function using labor as the only input, $Y_n = AN_n$, where A is an economy-wide productivity parameter and N_n is the amount of labor used in non-agriculture. Given prices, the representative firm in region u maximizes profits solving

$$\max_{N_{un}} P_{un}AN_n - W_uN_n, \quad (2)$$

where P_{un} and W_u are the price of the non-agricultural good and the wage per unit of labor in region u , respectively.

Farmers operating in the rural region use their fraction of land to produce agricultural goods in sector $s \in \{m, f\}$, according to the production function $y_s^i = A(z_s^i \ell)^{\alpha_s} (n_s^i)^{\beta_s} (x_j^i)^{\psi_s}$,

where n_s^i and x_s^i are hired labor and nonagricultural intermediate inputs, respectively, used by farmer i to produce crop s . I allow factor shares to be potentially different across agricultural goods and assume $\alpha_s + \beta_s + \psi_s = 1$. Since land is fixed for each farmer, there are decreasing returns at the farm level.

Given the choice to produce crop s , taking prices as given, a farmer maximizes profits by solving

$$\max_{\{n_s^i, x_s^i\}_{s \in \{m, f\}}} P_{rs} y_s^i - P_{rn} x_s^i - W_r n_s^i, \quad (3)$$

where P_{rs} is the price of good s and W_r is the labor wage in region r . Then, the payment received by each farmer is defined as $\pi_s^i = \alpha_s P_{rs} y_s^i$. The latter are residual earnings of a farm after input payments are made.

Interregional trade

Goods can be traded between regions subject to iceberg costs. Region j must ship τ_s^{jk} units of good s in order for one unit to arrive in region k . Thus, $\tau_s^{jk} = 1$ implies frictionless trade and $\tau_s^{jk} \rightarrow \infty$ implies autarky. By assumption, the rural region sends crops to the urban region and the latter sends non-agricultural goods to the rural region, so I omit the superscripts. Then, relative prices between regions are given by

$$\frac{P_{rn}}{P_{un}} = \tau_n, \quad \text{and} \quad \frac{P_{rs}}{P_{us}} = \frac{1}{\tau_s}, \quad s \in \{m, f\}. \quad (4)$$

Trade costs generate a wedge between prices across regions. In particular, trade costs increase the price of crops in the urban region and the price of intermediate inputs in the rural region. These trade technologies imply that interregional exports and imports, E_s and M_s , respectively, must satisfy the following restrictions

$$E_s = \tau_s M_s, \quad s \in \{n, m, f\}. \quad (5)$$

That is, trade costs increase the amount of goods that must be shipped to satisfy a given amount of demand in the destination region.⁷

3.4 Equilibrium

Farmers in the rural region choose to produce crop m or f based on their comparative advantage. In a competitive equilibrium, a farmer decides to produce maize if and only if her residual earnings of maize π_m^i are higher than her residual earnings of fruits π_f^i , that is, if and only if

$$\frac{z_m^i}{z_f^i} \geq \mathcal{K} \frac{(P_{rf})^{1/\alpha_f}}{(P_{rm})^{1/\alpha_m}} (W_r)^{\left(\frac{\beta_m}{\alpha_m} - \frac{\beta_f}{\alpha_f}\right)} (P_{rn})^{\left(\frac{\psi_m}{\alpha_m} - \frac{\psi_f}{\alpha_f}\right)}, \quad (6)$$

where $\mathcal{K}$ is a constant. Holding all else fixed, a lower relative price of fruits with respect to maize leads to a higher share of farmers producing maize. The direct effects of labor wages and price of non-agricultural inputs on the crop choice depends on how intensive is maize production in labor and intermediate inputs relative to fruits production. Below, I present a simplified case to illustrate how crop choices are affected by the key features of the model.

Additionally, the household of each region maximizes utility of its members (equation 1) by choosing consumption per capita of each good subject to income per capita I_j/N_j . Then, it can be shown that optimal consumption allocations in both regions are given by

$$\begin{aligned} c_{jm} &= \frac{\epsilon_m}{P_{jm}} \left(\frac{I_j}{N_j} - P_{jm} \bar{m} \right) + \bar{m}, \\ c_{jf} &= \frac{\epsilon_f}{P_{jf}} \left(\frac{I_j}{N_j} - P_{jm} \bar{m} \right), \\ c_{jn} &= \frac{\epsilon_n}{P_{jn}} \left(\frac{I_j}{N_j} - P_{jm} \bar{m} \right). \end{aligned} \quad (7)$$

Non-homothetic preferences imply that the expenditure share of maize decreases with income,

⁷Equations in 4 and 5 can be obtained from modeling the firm's maximization problem in a competitive transportation sector. See Herrendorf et al. (2012).

while the expenditure share of fruits and non-agricultural goods increases. These preferences are consistent with the patterns observed for budget shares of cereals and non-food products as income increases, and account for the fact that subsistence consumption is mostly observed for staple grains (see Appendix D). Household income in the urban region is given by labor payments, while income in the rural region is given by labor payments plus farmers earnings,

$$\begin{aligned} I_u &= W_u N_u, \\ I_r &= W_r N_{rw} + (N_r - N_{rw}) \left(\sum_{s \in \{m, f\}} \int_{i \in \Omega_s} \pi_s^i dG_i \right), \end{aligned} \quad (8)$$

where N_{rw} is the fraction of household members that are farmworkers in the rural region and Ω_s represents the set of farmers producing crop s .

The household problem means that every member receives the same utility within a region, so free movement of individuals implies that utilities are equalized; that is, individuals sort across regions until they are indifferent between living in either household. Moreover, the rural household decides the fraction of its members that operate farms and the fraction that are farmworkers. The first-order condition of the rural household with respect to the number of farmworkers implies that the return from hired labor must be equal to expected earnings of farmers,

$$W_r = \sum_{s \in \{m, f\}} \int_{i \in \Omega_s} \pi_s^i dG_i. \quad (9)$$

To define a competitive equilibrium, I assume the non-agricultural good is the numeraire and normalize $P_{un} = 1$. Then, market clearing conditions for goods are given by

$$N_u c_{un} + E_n = Y_n, \quad (10)$$

$$N_u c_{uf} = M_f, \quad (11)$$

$$N_u c_{um} = M_m, \quad (12)$$

$$N_r c_{rn} + X_{rm} + X_{rf} = M_n, \quad (13)$$

$$N_r c_{rf} + E_f = Y_f, \quad (14)$$

$$N_r c_{rm} + E_m = Y_m. \quad (15)$$

According to equations (10)-(12), total production of non-agricultural goods in the urban region is equal to local consumption by the household plus exports to the rural region, and urban consumption of crops is met by imports from the rural region. Equations (13)-(15) say that local consumption of non-agricultural goods in the rural region plus total intermediate inputs used by farmers, where $X_{rs} = (N_r - N_{rw}) \int_{i \in \Omega_s} x_s^i dG_i$, $s \in \{m, f\}$, are equal to imports from the urban region, and total production of crops, $Y_s = (N_r - N_{rw}) \int_{i \in \Omega_s} y_s^i dG_i$, $s \in \{m, f\}$, is equal to local consumption plus exports to the urban region.

Market clearing of labor market in the rural region requires that total labor demand equals the total number of farmworkers, $N_{rw} = N_{rm} + N_{rf}$, where $N_{rs} = (N_r - N_{rw}) \int_{i \in \Omega_s} n_s^i dG_i$. Note that in this model total labor in agriculture is equal to rural population: hired labor plus farm operators. That is, total labor allocated to crop s is given by $\mathcal{N}_{rs} = N_{rs} + \Phi_s$, where Φ_s is number of farmers producing crop s . The latter is consistent with the way in which total labor in crop production is calculated for the empirical evidence in Section 2.1.⁸ Finally, the fraction of land that every farmer receives satisfies $\sum_{s \in \{m, f\}} \Phi_s \ell = L$.

A *competitive equilibrium* with interregional trade is a set of prices of goods and inputs $\{P_{jn}, P_{jm}, P_{jf}, W_j\}$, $j \in \{u, r\}$; location choices (N_j individuals choose region j); farmers' earnings, π_s^i , $s \in \{m, f\}$; sets of households' allocations, $\{c_{jm}, c_{jf}, c_{jn}\}$, $j \in \{u, r\}$, and $\{\Phi_m, \Phi_f, N_{rw}\}$; a set of input choices in each region, $\{N_{un}, N_{rm}, N_{rf}, X_{rm}, X_{rf}\}$; and a set of interregional trade flows, $\{E_f, E_m, E_n, M_f, M_m, M_n\}$, such that: (i) given prices and farmers earnings, household allocations and individual location choices maximize utility in both regions; (ii) given prices, firms and farmers maximize profits; and (iii) market clearing conditions hold.

⁸According to Mexican national account data from 2013, non-hired labor (e.g. owners, family members and unpaid workers) account for 63 percent of total labor in agriculture.

3.5 Productivities distribution

I follow the parametrization of [Lagakos and Waugh \(2013\)](#) and define the joint distribution of crop-specific individual productivities as

$$G_j(z_m, z_f) = \mathcal{C}[F(z_m), F(z_f)], \quad \mathcal{C}[u, v] = \frac{-1}{\rho} \log \left(1 + \frac{(e^{-\rho u} - 1)(e^{-\rho v} - 1)}{e^{-\rho} - 1} \right). \quad (16)$$

$\mathcal{C}[F(z_m), F(z_f)]$ is a *Frank copula* with parameter $\rho \in \{-\infty, \infty\} \setminus \{0\}$. The latter governs the correlation between productivity draws, such that a positive value of ρ implies a positive dependence between z_m and z_f . The marginal distributions are Fréchet

$$F(z_s) = \exp(-z_s^{-\theta_s}),$$

where θ_s governs the dispersion of productivity draws and the scale parameter is normalized to one. There is a negative relationship between the value of θ_s and the variation of land augmenting productivity in crop $s \in \{m, f\}$; thus, a lower θ_s implies a higher variation in individual productivity. The dependence across productivity draws ρ and the variation of individual productivity θ_s determine the extent of alignment between absolute advantage and comparative advantage in a particular sector, that is, the difference in productivity between the marginal farmer and average farmer in a sector. The quantitative section provides more details on the role of these parameters in the model.

3.6 Trade costs and crop choice

Farmers select into crops as an efficient response to subsistence requirements of staple food and trade costs. To see how the model works, assume that $\alpha_m = \alpha_f$, $\gamma_m = \gamma_f$, and $\psi_m = \psi_f$. Then, the cutoff that determines the crop choice of farmers is given by

$$\frac{z_m^i}{z_f^i} \geq \left(\frac{P_{rf}}{P_{rm}} \right)^{1/\alpha} = \left[\left(\frac{\epsilon_f}{c_{uf}} \right) \frac{\tau_m}{\tau_f} \left(\frac{I_u/N_u}{P_{um}} - \bar{m} \right) \right]^{1/\alpha}. \quad (17)$$

In a world without trade costs, farmers would take prices of the urban market as given and decide which crop to produce based on the relative price; however, the existence of trade costs creates a wedge across regional prices and changes the relative price between crops in the rural region. Particularly, if farm-to-market trade costs are higher for fruits, then the relative price of fruits with respect to maize is lower and more farmers decide to produce maize.

Moreover, subsistence requirements raise the share of farmers producing maize by directly increasing the relative price of this crop. To see the interaction between trade costs and subsistence requirements in the model, note that when trade costs decrease in the economy there is a positive “income effect” that leads to an increase in the demand for non-agricultural goods and fruits that is higher than the increase in the demand for maize. The latter is due to the presence of non-homothetic preferences in the model. This implies that an economy with lower trade costs has a relatively lower demand for maize and, therefore, a smaller share of farmers producing this crop.

Finally, in this context trade costs can be considered a barrier that affects the allocation of labor across regions and types of crops. However, this is not a source of misallocation in the model as in [Hsieh and Klenow \(2009\)](#); to the contrary, farming decisions are efficient given the subsistence needs for staple food and the level of trade costs in each sector. The potential productivity gains from generating a movement of farmers from maize to fruit production come from the fact that a high concentration of farmers in maize production implies that many of them have a relative low productivity in that sector. In other words, farmers who have a higher productivity draw for fruit production might decide to produce maize because the relative price of this crop is high.

4 Quantitative analysis

In this section I calibrate the model to match features of the Mexican economy. In particular, I use farm data to discipline the distribution of land productivity in agricultural sectors and estimate interregional trade costs. Then, I introduce changes to the baseline economy to

evaluate the quantitative role of trade costs in allocating farmers across types of crops and generating low agricultural labor productivity. Specifically, I quantify the effects of assuming that there is no trade costs in the economy; however, while the latter case is helpful to analyze the overall importance of trade costs, it does not represent a plausible scenario for policy implications. Therefore, I also consider the counterfactual case of an overall decrease in trade costs to the U.S. level. Finally, to consider the general implications for poorer countries than Mexico, I recalibrate specific parameters of the model to match features of a typical African country and evaluate the effects of reducing trade costs in the economy.

4.1 Calibration

For the baseline case, I normalize the economy-wide productivity parameter A equal to one. Additionally, I follow the literature (e.g. [Restuccia et al., 2008](#); [Adamopoulos, 2011](#)) and set the total endowment of land L to match the land (harvested) to labor ratio in Mexico in 2014, which is equal to 0.42 hectares per worker. Thus, the remaining parameters that need to be calibrated are preferences weights ϵ_s , $s \in \{m, f, n\}$; the subsistence requirement of staple food $\bar{m}$; trade costs τ_s ; productivity distribution parameters θ_s and ρ ; and factor income shares $\{\alpha_s, \gamma_s, \psi_s\}$. To compare staples and cash crops, I focus on maize and the most important commercial fruits in Mexico: avocados, chili peppers, cucumber, melon, papaya, tomatoes, and watermelon.

Trade costs. The estimation of trade costs is based on the idea of comparing otherwise homogeneous products across origins and destinations in Mexico. To be specific, I compare prices of specific varieties of crops and chemical fertilizers. The assumption is that a crop variety is essentially the same good when it is sold by a farmer in a given state than when it is bought by a consumer in a wholesale market in another state. For example, in the case of fruits, if they did not spoil during transportation and are eatable by consumers, they are essentially the same good in the farm and the marketplace. Similarly, a particular type of fertilizer, such as Urea or Ammonium Sulfate, is a homogeneous chemical compound regardless of the point of sale.

Trade costs of fruits τ_f are estimated by comparing farm-gate prices and market prices of crops varieties between origins and destinations, as in equation (4) of the model. The origin price refers to the average farm price in each state obtained from ENA 2014 surveys and the destination price is the wholesale price in major cities of every state where it is sold (see Section 2.2). I focus on a subset of fruits varieties that are produced and/or sold in many states. These crops include avocado (hass), tomato (bola and saladette), watermelon (cambray), cucumber, papaya (maradol), and three varieties of chili pepper (poblano, jalapeño, and serrano). I aggregate trade costs of these fruits weighting each crop by its national production value in 2014. For staple crops, I estimate the mean regional price gap for maize (white).⁹ The results of this estimation are τ_f equal to 2.20 (45 percent farm share) and τ_m equal to 1.51 (66 percent farm share). These estimates imply large differences in trade costs within agriculture, which may reflect higher transportation and storage costs of fruits (e.g. refrigeration and spoilage), as well as monopoly power of intermediaries. The fact is that farmers face higher trade costs to enter fruits markets.

To estimate trade costs of non-agricultural inputs, I follow the steps described in Section 2.2, which imply comparing prices of fertilizers faced by farmers with market prices in the probable place of origin. I focus on those cases where the chemical fertilizer used is Urea, which is the most common fertilizer reported in the surveys. Given that most of this fertilizer is imported, I estimate the price gap between farms and the market price in Veracruz, one of the main ports where this product enters the country. The median price gap implies that τ_n is equal to 1.62, which is consistent with the fact that most farmers surveyed report high inputs costs as their main production obstacle.¹⁰

Jointly calibrated parameters. I jointly calibrate the subsistence parameter $\bar{m}$ and produc-

⁹The weights of fruits are: avocado hass 51 percent, tomato saladette 13 percent, chile jalapeño 11 percent, papaya maradol 9 percent, and the rest is distributed somewhat evenly among the other crops. White maize is the most produced variety of maize in the country. In 2014, white maize accounted for 90 percent of total production value of grain maize. Unlike yellow maize, which is mostly imported from the United States, white maize is traded domestically almost entirely.

¹⁰Other important port is Manzanillo located in Colima, but market prices of Urea in this state are practically the same as in Veracruz. I use the median gap because the distribution of relative prices of fertilizers has a long right tail and do not want to overestimate the size of trade costs for non-agricultural inputs.

Table 5: Model fit: targeted moments

Moment	Model	Data
Variance of log output value per hectare in fruits	2.82	2.82
Variance of log output value per hectare in maize	1.96	1.96
Ratio of avg. output value per hectare in maize to fruits	0.27	0.27
Ratio of total labor in maize to fruits	3.43	3.43

Notes: This table presents the results of the joint calibration in the model.

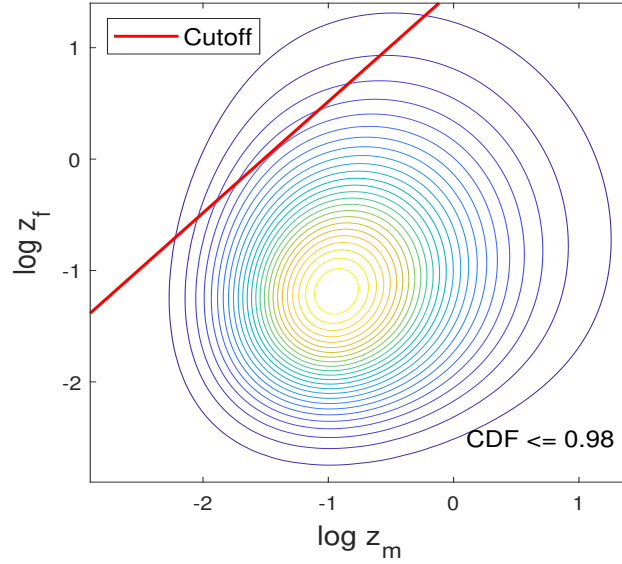
tivity distribution parameters θ_m , θ_f and ρ to match four moments estimated from the ENA 2014 surveys: total employment in maize relative to cash crops, 3.43; the ratio of average output value per hectare in maize to fruits, 0.27; and the variance of log yields in each agricultural sector, 2.82 for fruits and 1.96 for maize. The variance of yields represents the residual variation after controlling for farm size.¹¹ Table 5 shows that the model is able to match the targeted moments in the data.

The reasoning behind the joint calibration is the following. First, there is a positive relationship between the size of subsistence requirements of staple food and the share of total workers producing in that sector. Secondly, the variation in output value per hectare in each sector is governed by parameters θ_m and θ_f . To see this, note that in the model output value per hectare for farmer i in sector s is given by: $P_{rs}y_s^i/\ell = z_s^i \left(AP_{rs} \right)^{\frac{1}{\alpha_s}} \left(\frac{\beta_s}{W_r} \right)^{\frac{\beta_s}{\alpha_s}} \left(\frac{\psi_s}{P_{rn}} \right)^{\frac{\psi_s}{\alpha_s}}$; therefore, $\text{var}(\log(P_{rs}y_s^i/\ell)) = \text{var}(\log(z_s^i))$. This means that both parameters are disciplined by matching the conditional variance in the model with the observed variation in the data. Finally, the correlation parameter ρ governs the yield gap across types of crops by determining how strong is the relationship between absolute and comparative advantage.

The results presented in Table 6 imply that variation in fruits productivity is higher than in maize ($\theta_f < \theta_m$), which may reflect the fact that there is a wider variety of goods in the fruit sector, each of which requires particular farmer skills and/or land qualities to grow

¹¹The variance of log yields without controlling for the size of farms is 3.01 and 2.05 in fruits and maize, respectively. The quantitative results are not very sensitive to the difference between these targets.

Figure 7: Baseline probability distribution



Notes: This figure is obtained by simulating the probability distribution of land augmenting productivities $G(z_m, z_f)$ implied by the internal calibration of the model.

effectively. The positive value of ρ implies a Kendall rank correlation coefficient of 0.24; thus, the calibration implies a moderate positive correlation between productivity draws. Together, these results mean that farmers with relatively high productivity draws tend to select into fruit production. Note that in this case a very high correlation would mean that absolute and comparative advantage are not aligned in the maize sector because the marginal farmer would have a higher productivity than the average farmer in that sector (see [Lagakos and Waugh, 2013](#); [Young, 2014](#)).

To illustrate the results described in the previous paragraph, Figure 7 presents the probability distribution of individual productivity draws implied by the baseline calibration. This figure shows that farmers with high maize productivity choose this sector, but there is also a large fraction of farmers that produces maize even if their productivity is relatively higher in fruits. Moreover, the figure shows that farmers with low productivity draws self-select into staple crops, whereas fruit farmers are a more selective group with high productivity.

Preferences weights. Preferences weights govern expenditures shares as income tends to infinity and non-homothetic parameters become irrelevant. I follow the calibration strategy of [Gollin and Rogerson \(2014\)](#) and [Tombe \(2015\)](#), and use data from the 2005 International

Table 6: Calibration summary

Parameter	Value	Target
ϵ_m	0.09	Budget share for food, beverages and tobacco in rich countries
ϵ_f	0.11	Budget share for cereals relative to fruits in rich countries
β_m	0.37	Income share of labor in grains and oilseeds
ψ_m	0.23	Income share of non-agricultural inputs in grains and oilseeds
β_f	0.20	Income share of labor in fruits and vegetables
ψ_f	0.23	Income share of non-agricultural inputs in fruits and vegetables
τ_n	1.62	Interregional trade costs of fertilizers
τ_m	1.51	Interregional trade costs of maize
τ_f	2.20	Interregional trade costs of fruits
$\bar{m}$	0.08	Total labor in maize relative to fruits
θ_m	0.94	Variance of log output value per hectare in maize
θ_f	0.79	Variance of log output value per hectare in fruits
ρ	2.28	Ratio of average output value per hectare in maize to fruits

Comparison Program (ICP) to get budget shares for aggregates and food categories. I set these parameters to match the budget share for food, tobacco and beverages in rich countries, 0.20, and the budget share for cereals relative to fruits and vegetables in the same group of countries, 0.88. The latter implies that $\epsilon_m = 0.09$ and $\epsilon_f = 0.11$.¹²

Technology parameters. Factor income shares are calibrated using data from Mexico Input-Output tables for 2013. To estimate factor shares of maize in the model, I consider data of grains, legumes and oilseeds farming, while factor shares of fruits in the model are estimated using data of fruits, nuts and vegetables farming. Payments to labor are calculated adjusting compensation to employees following Gollin (2002), that is, I impute the employee compensation of non-hired labor (owners, family, contract labor, and non-remunerated labor). For each

¹²To put these numbers in context, low-income countries spend 49 percent of their budget on food and the share for cereals is 1.29 times greater than the share for fruits and vegetables.

Table 7: Quantitative implications

Non-targeted Moment	Model	Data
Agriculture Employment Share (percent)	14.82	13.77
Ratio of VA per Worker in Maize to Fruits	0.36	0.41

Notes: Aggregate productivity gaps are reported. Labor productivity is equal to total value added divided by total labor (hired labor plus non-hired) in each sector. Labor productivity is estimated using data from ENA 2014 and employment in agriculture for 2014 is obtained from the World Bank.

sector, I calculate the average compensation of remunerated workers and multiply it by total workers (hired and non-hired). The share of non-agricultural intermediate inputs is computed using expenditures on inputs from non-agricultural sectors. Finally, I assume that payments to land include farm profits, so these payments are estimated as gross operating surplus minus the compensation of non-hired labor. The latter adjustment is made so gross value added in the industry remains unchanged. Results are reported in Table 6. According to these estimates, the income share of labor is higher in grains production than in fruits. This is in line with labor intensive techniques used in developing countries for subsistence agriculture.

Next, I validate the model by looking at other quantitative implications. Table 7 compares the baseline results with relevant non-targeted moments in the data. The model is able to replicate a share of agricultural employment that is very similar to the one observed in Mexico. Moreover, the model matches well the labor productivity gap between types of crops. This means that in the baseline economy both labor and land productivity are higher in fruits than in maize farming. The latter is especially important since these productivity gaps are the main empirical motivation of the paper. A model without heterogeneous farmers could not replicate these results since differences in productivity across agricultural sectors would only reflect differences in income shares of inputs. In such case, the productivity gaps could only be generated if maize production is significantly more intensive in both land and labor compared to fruits, or if explicit barriers or wedges are introduced to prevent the equalization of marginal products across sectors.

Additionally, the value of $\bar{m}$ in the model represents 23 percent of total maize production

and according to Mexican government data 18 percent of white maize production is used for subsistence consumption. This share is similar for beans, which are another important staple in Mexico. In comparison, subsistence consumption of grains in poorer Central American countries accounts for nearly 50 percent of production ([SAGARPA, 2017](#); [FAO-RUTA, 2010](#)). Furthermore, the model matches the lower relative price of fruits with respect to maize in rural regions. This is a direct implication of the calibration because trade costs are estimated based on regional price gaps: $\frac{P_{rf}}{P_{rm}} \bigg/ \frac{P_{uf}}{P_{um}} = \frac{\tau_m}{\tau_f} < 1$. A similar argument applies to the relative price of intermediate inputs (fertilizers) with respect to crops between regions. Lastly, the equilibrium value of land per farm ℓ is equal to 4.6, which is close to the median farm size in Mexico (approx. 3 hectares) and consistent with the fact that almost 70 percent of farms have at most 5 hectares.

4.2 Quantitative Experiments

In this section I carry out multiple counterfactual experiments. First, I assess the impact of assuming that there is no trade costs in the economy, that is, τ_s equal to one for every sector. This case is useful to analyze the overall importance of trade costs. Nevertheless, since the latter is not a plausible scenario for policy implications, I consider a benchmark that is consistent with equivalent trade costs in the United States. Furthermore, I present experiments using alternative model specifications to highlight the role of different assumptions or mechanisms in the baseline model.

To assess the quantitative importance of trade costs, I focus on the allocation of labor and land across crops, agricultural value added per worker, and the share of employment in agriculture. I also quantify the effects on the amount of modern inputs per worker used in agricultural production. Lastly, I use a Fisher price index in the rural region to compare agricultural value added in the baseline economy and the counterfactual cases.

The results from assuming zero trade costs across regions are presented in the second column of Table 8. Agricultural labor productivity increases by 21 percent, the ratio of total employment in fruits to maize increases by 17 percent, and there is similar reallocation of land

Table 8: Counterfactuals: reducing trade costs

Percentage change, relative to baseline	No Trade Costs ($\tau = 1$)	U.S. benchmark (55 percent reduction)
Agricultural Value Added per Worker	20.5	13.4
Fruits to Maize Labor Ratio	16.7	14.6
Fruits to Maize Land Ratio	17.6	15.3
Intermediates to Labor Ratio (Agriculture)	27.6	18.2
Agricultural Employment (pp change)	2.1	1.6

Notes: This table presents the percentage change with respect to the baseline values of reducing trade costs in the model. For agricultural employment, the change refers to percentage points. Cobb-Douglas technologies imply that changes in intermediates to labor ratio are the same for both crops.

across crops (which is equivalent to a reallocation of farmers in the model). These results imply that eliminating trade costs leads to an improved allocation of farmers across types of crops based on comparative advantage. Additionally, the use of intermediate inputs relative to labor increases by 28 percent, which has a positive effect on agricultural labor productivity. These results have important implications for the way we think about structural transformation. That is, productivity gains in the economy, such as improvements in transportation technologies, can generate reallocation of labor between agricultural sectors and not only between agriculture and non-agriculture.

Note that the population share in agriculture increases by two percentage points, from 14.8 percent of total employment to 16.9 percent. On one hand, more people can move to the city because food is less costly to transport and more intermediate inputs are used in agriculture. On the other hand, enough labor needs to work in fruits farming to satisfy the relatively higher demand (income effect) and individuals do not need to live in the city to consume non-agricultural goods at a lower price. In this case, the second force is moderately stronger so rural population increases. Similar effects on agricultural employment are found by [Herrendorf et al. \(2012\)](#) and [Gollin and Rogerson \(2014\)](#) when trade costs decrease in the economy.

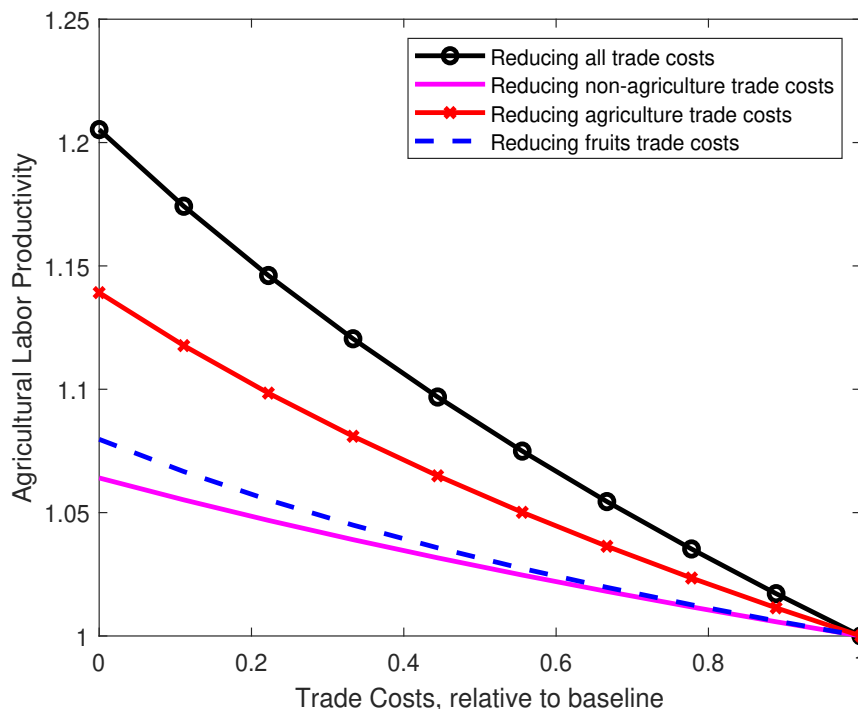
Next, I take the United States as a low trade costs benchmark. Price comparisons across regions in Mexico were based on farm-gate prices and prices in wholesale markets. To make the equivalent calculation for the United States, I compare the farm share of total retail costs in fruits markets with the accumulated cost share of farms, transportation, and wholesale trade using data from the USDA in 2007. The latter implies a farm-price share of 65 percent, which means that trade costs of fruits in the United States are around 55 percent lower than Mexico.¹³ Using this benchmark, I reduce all trade costs by the same proportion. The idea of this experiment would be an improvement in the quality of transportation and storage facilities in Mexico, or the adoption of policies inducing competition in transportation markets, that would reduce trade costs to the U.S. level. I focus on fruits due to data availability, however, these goods are the most sensitive to transportation costs.

The magnitude of the results presented in the third column of Table 8 is fairly large. Agricultural labor productivity increases by 13 percent and the ratio of employment in fruits to maize increases by 15 percent; furthermore, the lower costs of modern inputs increase the intensity with which they are used in the production of agricultural goods by 18 percent. Therefore, reducing trade costs to the U.S. level would raise agricultural labor productivity in Mexico by allocating more farmers in high-productivity cash crops and increasing the relative amount of modern inputs used in agricultural production. The size of the productivity gains are low compare to the large agricultural productivity gaps in Gollin et al. (2014a); however, these results are in line with other papers looking at the effects of transportation improvements on agricultural productivity. For example, Sotelo (2019) finds an increase of 5 percent in agricultural productivity from paving roads in Peru, and Donaldson (2018) finds that railroad access increased real agricultural income by 16 percent in colonial India.

To complement the previous experiments, Figure 8 shows the relationship between reductions in trade costs and agricultural labor productivity. In addition to decreasing trade costs in every sector, I reduce trade costs in each sector independently. These results suggest that

¹³A farm share of 65 percent implies that $\tau_f(\text{US})=1.54$. Then, the ratio of trade costs in the United States to Mexico can be calculated as $(\tau_f(\text{US}) - 1)/(\tau_f(\text{MX}) - 1)=0.45$. U.S. data from www.ers.usda.gov/data-products/food-dollar-series/food-dollar-application.aspx.

Figure 8: Productivity gains from reducing trade costs



Notes: Agricultural labor productivity refers to value added per worker in agriculture relative to the baseline economy. Trade costs in other sectors are held constant in every case.

meaningful improvements in transportation costs are needed in order to obtain significant gains in agricultural labor productivity. The latter provides support to the large amounts of resources that developing countries and international organizations allocate to improve transport infrastructure. Moreover, the results show that reducing trade costs of crops has a relatively large effect on agricultural productivity, whereas reducing trade costs of modern inputs, keeping trade costs of crops constant, has a smaller effect on agricultural productivity.

To measure the welfare gains of these experiments, I obtain the amount of income that would make the household of each region indifferent between the baseline case and the counterfactual economy, and calculate the average (population-weighted) of both regions. The quantitative results imply that there are large welfare gains from eliminating trade costs in the economy: 20 percent in case of no trade costs and 13 percent in the U.S. benchmark.

Given that transportation technology is modeled as an iceberg cost, it is important to distinguish how much of welfare gains in the model are due to the lower spoilage that results

Table 9: Welfare gains decomposition: zero trade costs ($\tau_s = 1$)

	Lower Spoilage	General equilibrium
Agregate welfare gains (percent)	55.36	44.64

Notes: The welfare gains from lower spoilage are calculated by assuming that all spoilage gains are consumed by the destination region in each sector, keeping all choices fixed. The difference between these and the total gains are the general equilibrium gains.

from decreasing these costs and how much is due to general equilibrium effects. To measure the former, I keep every decision of the baseline economy fixed and increase consumption of agricultural and non-agricultural goods in the urban and rural regions by the change in imports when trade costs decrease. That is, I assume households consume the extra amount of goods that they receive without altering their baseline decisions. The general equilibrium gains are then the total welfare gains minus the welfare gains from lower spoilage in the economy. The decomposition from a counterfactual with zero trade costs is presented in Table 9. The general equilibrium gains are large and represent almost half of the total gains. Thus, the gains from reducing trade costs are not only due to the lower spoilage of goods, but also to the fact that agents optimally react to trade improvements by reallocating resources across sectors, in particular, across crops within agriculture.

Next, I present additional experiments to highlight the role of different assumptions and mechanisms in the model. To be specific, I look at the effects of reducing trade costs on equilibrium allocations using a different calibration or model specification. For each of the experiments presented below, I repeat the joint calibration to match the targeted moments in the data.

In the first experiment, I analyze the importance of having different labor shares in the production function of crops by setting the labor share of fruits equal to the one of maize. The results are presented in the third column of Table 10 and have the following implications. First, the change in labor and land allocation is much larger compared to the baseline. One reason is that now input intensity is equal across crops, so the need for labor in maize production is relatively less important in this case. Also, the calibrated value of $\bar{m}$ needs to be higher in

Table 10: Reducing trade costs ($\tau_s = 1$) with alternative specifications

Percentage change	Baseline	Same labor shares $\beta_f = \beta_m$	No subsistence $\bar{m} = 0$	No intermediates $\psi_s = 0$
Fruits to Maize Labor Ratio	16.7	37.8	0.0	14.3
Fruits to Maize Land Ratio	17.6	44.9	0.0	15.1
Ag. Employment (pp change)	2.1	0.1	3.1	2.6

Notes: This table compares the results of the baseline case with alternative model specifications. See text for details.

order to match the employment gap across crops, which means that non-homothetic effects are stronger. Another implication of the latter is that the change in agricultural employment is smaller in the alternative case.

In the second experiment, I set subsistence requirements of staple food $\bar{m}$ equal to zero. The results are presented in the fourth column of Table 10. The main lesson from this experiment is that without non-homothetic preferences there is no reallocation of labor across crops because labor shares in the economy are fixed based on constant expenditure shares. In the baseline model, trade costs amplify the selection of farmers into maize because there is a subsistence requirement of staple food. Furthermore, the change in agricultural employment is larger compared to the baseline. This is because individuals are pulled to the rural region by lower non-agricultural prices and without subsistence requirements the income elasticity of demand for all goods is the same, so there are no offsetting effects.

Finally, in a third experiment, I assume there are no intermediate inputs in agricultural production by setting ψ_s equal zero. The last column of Table 10 shows that in this case reallocation of labor across crops is smaller. This is because the presence of intermediate inputs generates an additional negative income effect that pushes workers to staple crops when such intermediates are costly to acquire. The calibrated value of $\bar{m}$ needs to be higher in the alternative case, which partially offsets the previous effect. Intuitively, in a model with trade costs and intermediate inputs, more labor needs to be allocated to grow staple crops in order to produce enough food for the population.

To summarize, the staple productivity puzzle documented in Section 2.1 is generated in the model because a large share of low-productivity farmers selects into staple crops, while fruit producers are a selective group with relatively high productivity. Furthermore, the counterfactual results imply that the interaction of subsistence requirements and interregional trade costs can partially account for the puzzling allocation of labor across types of crops in Mexico. Lastly, these results suggest a different approach to structural transformation by linking productivity or technological improvements to reallocation of labor from low to high productivity sectors within agriculture.

4.3 Application to Uganda

The motivation of this paper was also based on the fact that many poor countries allocate most of their land to low productivity staple crops, even if yields in many fruits are significantly higher (see Figure 3). Then, to analyze the quantitative implications of the model for poorer countries than Mexico, I recalibrate the baseline economy to match features of a typical African country: Uganda. In particular, I calibrate the economy-wide productivity parameter and the trade costs to match the share of employment in agriculture and the price gap of fruits across distant regions in that country. The idea is to change the baseline economy in order to get a higher share of employment in agriculture, as in a typical poor country, and be consistent with the low quality of transportation infrastructure.

According to United Nations data from 2014, agriculture accounted for 72 percent of employment in Uganda. Then, I decrease A to make the modeled economy poor enough so that more people work in agriculture. The latter is a result of non-homothetic preferences in the model. In addition, [Gollin and Rogerson \(2010\)](#) compute the difference between the wholesale price of Matoke (a variety of banana) at the region of origin and the wholesale price at distant points of sale. The highest ratio of destination to origin price is 4.17, with a distance between points of approximately 500 kilometers (311 miles). The latter implies that fruits trade costs in Uganda are 2.6 times higher than in Mexico. That regional price gap represents the cost of transporting fruits from the southwest region to the north region of Uganda, and

Table 11: Reducing trade costs in Uganda

Percentage change, relative to baseline	U.S. benchmark (83 percent reduction)
Agricultural Value Added per Worker	209.1
Fruits to Maize Labor Ratio	330.7
Fruits to Maize Land Ratio	274.4
Intermediates to Labor Ratio (Agriculture)	409.9
Agricultural Employment (pp change)	-39.9

Notes: This table presents the percentage change of reducing trade costs. The baseline refers to the calibrated model for Uganda. For agricultural employment, the change refers to percentage points. Cobb-Douglas technologies imply that changes in intermediates to labor ratio are the same for both crops.

thus reflects the quality of transport infrastructure across the country. Then, similar to the experiment based on the United States, I use this number to increase all trade costs in the economy proportionally.

Once the baseline economy has been calibrated to match the facts described in the previous paragraph, I reduce trade costs to the U.S. level and quantify the effects as was done for the case of Mexico. In this case, trade costs in Uganda are reduced 83 percent to reach the U.S. level.¹⁴ The results presented in Table 11 show that the effects are much larger for a poor country like Uganda. Agricultural labor productivity increases by more than a factor of 3, the ratio of employment in fruits to maize is over 4 times larger, and the use of intermediates relative to labor increases by a factor of 5. The latter reflects the lower cost of modern inputs and the large movement of population from the rural region to the city. These results are suggestive that reducing trade costs in extremely poor countries can increase agricultural labor productivity by releasing individuals from this sector and allocating a larger share of agricultural labor to high-productivity crops. The fact that gains from trade are especially large for poor countries is a common result in the trade literature (see e.g. Adamopoulos, 2011 and Tombe, 2015).

¹⁴The magnitude of the difference in trade costs between Uganda and the United States is consistent with the findings of Adamopoulos (2011) for differences in transportation costs between rich and poor countries.

5 Discussion

The quantitative results of model imply that trade costs are an important factor affecting labor and land allocations across types of crops. However, the model abstracts from other factors that could also be important for crop choices on their own or by interacting with interregional trade costs. In this section, I discuss some of the simplifications in the model and their potential implications.

First, the model considers a closed economy without international trade. In the case of Mexican agriculture, international trade is important. Mexico is a net importer of grains and a net exporter of fruits. To be specific, according to 2013 Input-Output Tables, net imports were equal to 43 percent of GDP in maize farming, and net exports represented 49 percent of GDP in fruits and vegetables farming. That said, there is a crucial caveat with those numbers. Most of the maize that Mexico imports is yellow corn that is mostly used for animal feed (76 percent) and just a very small share is used for subsistence and human consumption (4 percent). In contrast, Mexico is self-sufficient in production of white maize and total imports represent less than 5 percent of total production; moreover, around 70 percent of white maize is used for subsistence and human consumption ([SAGARPA, 2017](#)). Thus, the definition of staple crop in the closed economy model is consistent with the situation of white maize in Mexico. This variety of maize accounted for 90 percent of maize production (as grain) in 2014 and it is traded domestically almost entirely.

Now, abstracting from international trade seems more important for fruits given the large size of exports. However, while it is true that international demand is a key determinant of fruits production, the fact that productivity is significantly larger in fruits farming, yet most labor is allocated to maize, is still puzzling unless one assumes that domestic trade cannot increase anymore. A model with international trade would certainly include the possibility to intensify the specialization in fruits farming and increase the imports of maize from more productive countries. In that respect, the results from the closed economy model could be considered a *lower* bound of a model featuring international trade. If reducing interregional

trade costs improves the opportunities to export fruits and import maize, then the reallocation of labor from maize to fruits would be even larger.

Another factor that the model does not consider are land markets. This assumption seems innocuous for the case of Mexico where the *Ejido*, a system of communal land, limits the trade of land in spite of reforms done in the 1990s.¹⁵ Furthermore, large commercial farms in Mexico are not only fruits producers, in fact, some of the main commercial producers of maize in the country are large farms located in the northwest of the country. Thus, increasing the operated land size seems crucial for transition to commercial farming in either type of crop, but the connection is less clear for the choice between maize and fruits by small and medium scale farmers.

It is worth mentioning that Table 2 showed that controlling for farm size reduces the productivity gaps across types of crops, but still leaves sizable difference in labor and land productivity between fruits and maize. That said, land misallocation within agriculture has shown to be important in explaining low agricultural productivity in poor countries (e.g. Adamopoulos and Restuccia, 2014; Restuccia and Santaaulalia-Llopis, 2017). In a model with farm-size distortions, allowing small farms to operate more land might be important to achieve the proper scale to participate in crops markets that are export or commercially oriented.

Finally, the literature on agricultural productivity gaps has also considered the importance of risk for farming decisions (see Donovan, 2018). At first glance, it seems like risk could be an important determinant of crop choices given that most fruits have higher returns than maize, but their variance is also higher. However, the results presented in Appendix C using historical data at the state level, show that the yield distribution in fruits farming practically first-order stochastically dominates the yield distribution in maize farming. The difference in average yields is so large that it does not seem plausible that risk is the main barrier affecting crop choices. Having said that, a factor that might be important for crop choices in poor countries are the initial costs of switching from maize to fruits farming, such as buying trees or proper

¹⁵According to the 2007 agricultural census, 68 percent of farms (units of production) were part of Ejidos and they had 34 percent of the land.

seeds. Farmers might not have the capital that is needed to cover those costs. Moreover, in the presence of large adjustment costs, risk might actually matter for crop choices.

6 Conclusions

This paper documents evidence that labor productivity in agriculture is much lower for staple crops than for cash crops. I use microdata from Mexican farms to show that many fruits have a higher labor productivity than maize, yet the share of employment in the latter is significantly larger. These findings imply that the agricultural productivity puzzle is largely a staple productivity puzzle, so focusing on production decisions of farmers is key to understand why agricultural labor productivity is so low in poor countries.

One explanation proposed in this paper is that a high share of farmers decides to produce staple crops due to subsistence requirements of staple food and the existence of interregional trade costs in agricultural markets. If trade costs in Mexico were at the U.S. level, value added per worker in agriculture would be 13 percent higher. This productivity gain is driven by the reallocation of labor from staple to cash crops and by a higher use of modern inputs in agriculture. These results can be related to the findings of [Lagakos and Waugh \(2013\)](#) by thinking of high interregional trade costs as one component of low productivity in poor countries, with a particular effect of increasing the allocation of labor in low-productivity staple crops.

Moreover, the findings of this paper suggest a different approach to structural transformation, namely, that sectoral productivity gains or technology improvements could also reallocate labor within agricultural sectors. Similar to the decomposition of services studied by [Duarte and Restuccia \(2019\)](#), the results of this paper imply that we can also decompose changes in agriculture. Indeed, based on the case of the United States where maize production is highly mechanized, a characteristic of the structural transformation within agriculture could be a higher allocation of agricultural employment to labor intensive crops such as fruits.

There are alternative and complementary explanations to the one proposed in this paper.

For example, switching from staple to cash crops might require large initial investments such as buying a fruit tree or acquiring modern seeds to grow attractive commercial crops; thus, barriers preventing access to capital and input markets may keep too many farmers out of the cash crops sector. Also, farmers need to have accurate and updated information on crop prices in order to make the best farming decision; so barriers to the flow of information might be key to explain why many farmers decide to grow maize. These and other possible explanations are subject of future research.

Lastly, the results of this paper have important policy implications. First, reducing storage and transportation costs of crops can have significant positive results on agricultural labor productivity. Second, policies should focus on guaranteeing competitive conditions along the supply chain in agricultural markets. Reducing transaction costs and establishing competitive markets seems crucial to allow farmers to enter and grow in profitable agricultural markets.

References

- Adamopoulos, Tasso.** 2011. “Transportation Costs, Agricultural Productivity, and Cross-Country Income Differences.” *International Economic Review*, 52 (2): 489–521.
- Adamopoulos, Tasso, and Diego Restuccia.** 2014. “The Size Distribution of Farms and International Productivity Differences.” *American Economic Review*, 104 (6): 1667–1697.
- Adamopoulos, Tasso, and Diego Restuccia.** 2018. “Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data.” Working Paper.
- Adamopoulos, Tasso, and Diego Restuccia.** 2020. “Land Reform and Productivity: A Quantitative Analysis with Micro Data.” *American Economic Journal: Macroeconomics*, Forthcoming.
- Alder, Simon.** 2019. “Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development.” Working Paper.
- Caselli, Francesco.** 2005. “Accounting for Cross-Country Income Differences.” *Handbook of Economic Growth*, edited by Philippe Aghion and Steven Durlauf, Elsevier Science: 679–741.
- Caunedo, Julieta, and Elisa Keller.** 2019. “Capital Obsolescence and Agricultural Productivity.” Working paper.
- Chen, Chaoran.** 2017. “Untitled Land, Occupational Choice, and Agricultural Productivity.” *American Economic Journal: Macroeconomics*, 9(4): 91–121.
- Chen, Chaoran.** 2020. “Technology Adoption, Capital Deepening, and International Productivity Differences.” *Journal of Development Economics*, 143.
- Donaldson, Dave.** 2018. “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure.” *American Economic Review*, 108 (4-5): 899–934.

- Donaldson, Dave, and Richard Hornbeck.** 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *Quarterly Journal of Economics*, 131 (2): 799–858.
- Donovan, Kevin.** 2018. "Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences." Working paper.
- Duarte, Margarida, and Diego Restuccia.** 2019. "Relative Prices and Sectoral Productivity." *Journal of the European Economic Association*, Forthcoming.
- Eaton, Jonathan, and Samuel Kortum.** 2002. "Technology, Geography, and Trade." *Econometrica*, 70 (5): 1741–1779.
- FAO-RUTA.** 2010. "Small producers of basic grains in Central America." PESA Centroamérica y Unidad Regional de Asistencia Técnica (RUTA).
- Gollin, Douglas.** 2002. "Getting Income Shares Right." *Journal of Political Economy*, 110 (2): 458–474.
- Gollin, Douglas, and Richard Rogerson.** 2010. "Agriculture, Roads, and Economic Development in Uganda." NBER Working Paper No. 15863.
- Gollin, Douglas, and Richard Rogerson.** 2014. "Productivity, Transport Costs and Subsistence Agriculture." *Journal of Development Economics*, 107: 38–48.
- Gollin, Douglas, David Lagakos, and Michael E. Waugh.** 2014a. "The Agricultural Productivity Gap." *The Quarterly Journal of Economics*, 129 (2): 939–993.
- Gollin, Douglas, David Lagakos, and Michael E. Waugh.** 2014b. "Agricultural Productivity Differences across Countries." *American Economic Review: Papers and Proceedings*, 104 (5): 165–170.
- Gottlieb, Charles, and Jan Grobovšek.** 2019. "Communal Land and Agricultural Productivity." *Journal of Development Economics*, 138: 135 – 152.

- Herrendorf, Berthold, and Todd Schoellman.** 2015. “Why is Measured Productivity so Low in Agriculture?” *Review of Economic Dynamics*, 18 (4): 1003–1022.
- Herrendorf, Berthold, James A. Schmitz, and Arilton Teixeira.** 2012. “The Role of Transportation in U.S. Economic Development: 1840-1860.” *International Economic Review*, 53 (3): 693–715.
- Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi.** 2014. “Growth and Structural Transformation.” In *Handbook of Economic Growth*. Vol. 2 of *Handbook of Economic Growth*, Chapter 6, 855–941. Elsevier.
- Hsieh, Chang-Tai, and Peter J. Klenow.** 2009. “Misallocation and Manufacturing TFP in China and India.” *Quarterly Journal of Economics*, 124 (4): 1403–1448.
- Lagakos, David, and Michael E. Waugh.** 2013. “Selection, Agriculture, and Cross-Country Productivity Differences.” *American Economic Review*, 103 (2): 948–980.
- Le, Kien.** 2020. “Land Use Restrictions, Misallocation in Agriculture, and Aggregate Productivity in Vietnam.” *Journal of Development Economics*, 145.
- OECD.** 2007. “Agricultural and Fisheries Policies in Mexico.” Organization for Economic Cooperation and Development.
- Restuccia, Diego, and Raul Santaeulalia-Llopis.** 2017. “Land Misallocation and Productivity.” Working Paper.
- Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu.** 2008. “Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis.” *Journal of Monetary Economics*, 55: 234–250.
- SAGARPA.** 2017. “Agricultural National Plan 2017-2030: Yellow and White Mexican Maize Grain.” Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, Mexico.

- Sotelo, Sebastian.** 2019. “Domestic Trade Frictions and Agriculture.” *Journal of Political Economy*, Accepted.
- Tombe, Trevor.** 2015. “The Missing Food Problem: Trade, Agriculture, and International Productivity Differences.” *American Economic Journal: Macroeconomics*, 7 (3): 226–258.
- Vollrath, Dietrich.** 2009. “How Important are Dual Economy Effects for Aggregate Productivity?” *Journal of Development Economics*, 88: 325–334.
- Young, Alwyn.** 2014. “Structural Transformation, the Mismeasurement of Productivity Growth, and the Cost Disease of Services.” *American Economic Review*, 104 (11): 3635–3667.

Appendix

A Data and empirical details

This section provides more details on the evidence presented in Section 2. Using data from the ENA 2014 surveys, the steps to calculate value added for each crop that is produced in a farm are the following. First, to obtain the value of output I multiply the volume of harvested output by the farm-gate price reported by the farmer. Many farmers do not report a price because they did not sell any output in that period of time, especially those who produce maize; for such cases, I use the average price of the crop in the municipality where the farm is located or, in cases where there is no data to compute the latter, the average price of the crop in the state. I eliminate outliers (0.5 percent of each tail) to compute those average prices. The next step is obtaining the value of intermediate inputs used in the production of each crop. Different categories of farm expenses are reported at the crop level; however, some categories like soil preparation and sowing may include payments to capital and labor. Therefore, I only consider expenditures on modern inputs that do not include any payments to factors of production: fertilizers; chemicals and pesticides; and irrigation. Finally, I subtract the value of production that is used for seed and feed in the farm from the value of total output.

To estimate value added per hectare, I take the amount of harvested hectares reported by the farmer for each crop. Estimating value added per worker involves additional steps. The number of workers (owner, family members and hired labor) are reported at the farm level and, thus, they might be used in the production of more than one crop within a farm. Then, I focus on farms that only produce one type of crop, otherwise there is not an obvious way to allocate labor to different crops produced in a farm. To define farms producing a single crop, I aggregate the different varieties of the crops considered in Section 2.2 into one category; for example, all varieties of chili pepper are considered as one type of crop. Under such considerations, from the total number of farms producing maize, fruits or both, only 5.8 percent of them produces maize and one of the fruits. Thus, by focusing on farms that produce

one type of crop I only lose a small share of farms.¹⁶ However, the farms omitted from the estimation of labor productivity might grow any number of different crop varieties and that is reflected in the difference of farm-crop observations between land and labor productivity gaps reported in Table 2.

I now describe the farm price data used in the estimation of trade costs. For crops, I use prices reported by farmers and compute the average price at the state level after eliminating outliers (0.5 percent of each tail). In the case of fertilizers, farmers do not report the actual price they paid, instead they report total quantities of both natural and chemical fertilizers used for crop production, as well as total expenses on fertilizers. Since I cannot split the latter between natural and chemical fertilizers, I compute the fertilizer price as total expenditure divided by total quantity. This procedure results in a distribution of fertilizers prices with fat tails. From observation and comparison with public market prices, I eliminate outliers to compute average prices at the state and municipality level (2.5 percent of each tail). In Section 2.2 of the main text, I focus on cases where the chemical fertilizer used by the farmer was Urea or Ammonium Sulfate, which account for 45 percent of the observations that reported fertilizers.

As described in the main text, the second most important data source used for the empirical evidence comes from the SNIIM. This is a government website that provides information of market prices in primary sectors of the economy. I build a dataset with the monthly price of varieties of fruits and grains in the main wholesale markets of the country (usually located in the capital of a state). For each crop in a particular market, both the price (per kilogram most of the times) and state of origin are reported. It is worth mentioning that downloading and processing these data requires a non-trivial amount of time in order to produce a database of market prices of crop varieties in every state. Additionally, SNIIM provides data on market prices of fertilizers throughout the country. I use this to obtain the price of Urea and Ammonium Sulfate in the possible state of origin, either a port or the state where a production plant is located.

¹⁶In general, the surveyed farms produce less than two (1.7) different varieties of crops on average.

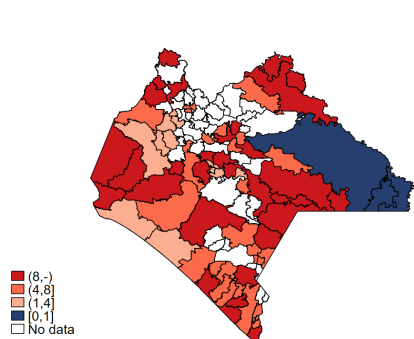
Finally, I use public data from national accounts. Particularly, I use the 2008 and 2013 Input-Output matrices that have data on value added, inputs expenditures, and total employment at the six-digit industry level, including 50 types of crops. I also use agricultural data aggregated to the municipality and state level from the *Sistema de Información Pesquera y Alimentaria* (SIAP) of the *Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación* (SAGARPA). These data on production, prices, yields, and land is collected by governments offices located in many localities throughout the country and is available from 1980 to 2014.

B Regional production of crops

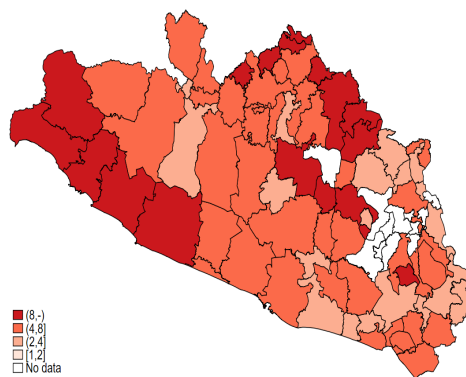
This section provides additional evidence that farmers in many regions of Mexico have the possibility to produce cash crops with a higher productivity than maize, so farmers are not always forced to grow staple crops by conditions related to climate or quality of soil in a particular area. I look at the geographical distribution of yields, measured as output value per hectare, within states in Mexico using data from SIAP. To simplify the analysis, I focus on a particular group of crops that includes the types of fruits considered in the main text: avocado, banana, chili pepper, cucumber, mango, papaya, tomato, and watermelon; and on those states with a high level of poverty and large share of agricultural employment. Every state is divided in municipalities that cover multiple towns or cities. There are 2,457 municipalities in the country and 53 percent are located in the six states considered.

Figure 9 presents the distribution of yields across municipalities in 2014. These maps show that most municipalities produce fruits with yields that are substantially higher than maize. Only Chiapas and Oaxaca have relatively big municipalities with no data, but these are surrounded by localities with presence of high productivity fruits. Thus, even if some subregions have higher levels of productivity, there is no evidence to conclude that maize is the only feasible option in most places.

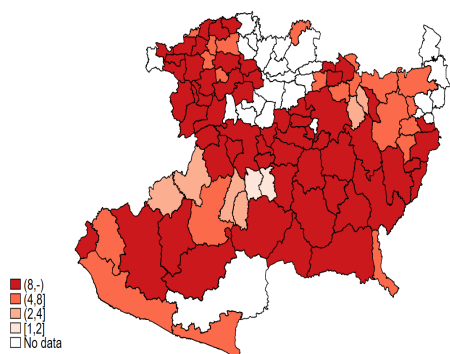
Figure 9: Distribution of relative yields of fruits in poor states
(Maize=1)



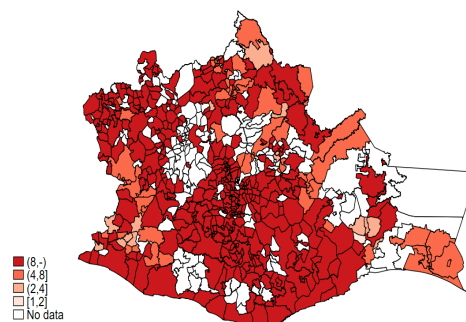
Chiapas



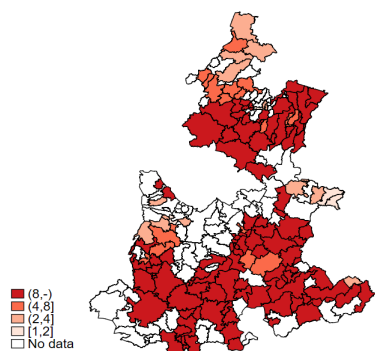
Guerrero



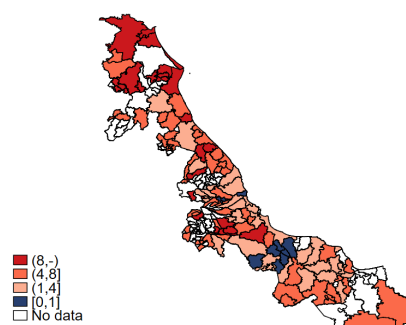
Michoacan



Oaxaca



Puebla



Veracruz

Note: Average yields are weighted by harvested hectares.
Source: Author's estimates using SIAP data, 2014

Table 12: Productivity gaps and employment shares crops

Crops	Value added per worker relative to non-agriculture	Employment share of total agriculture (percent)
Maize	0.15	23.1
Other grains	0.19	15.8
Top cash crops	0.40	8.2
Agricultural Sector	0.16	100

Notes: Top cash crops includes avocados, tomatoes, chili peppers, other vegetables, and other non-citrus fruits and nuts. These crops account for 64 percent of total exports in agriculture and 54 percent of Value Added in fruits and vegetables farming. Products are classified according to the NAICS.

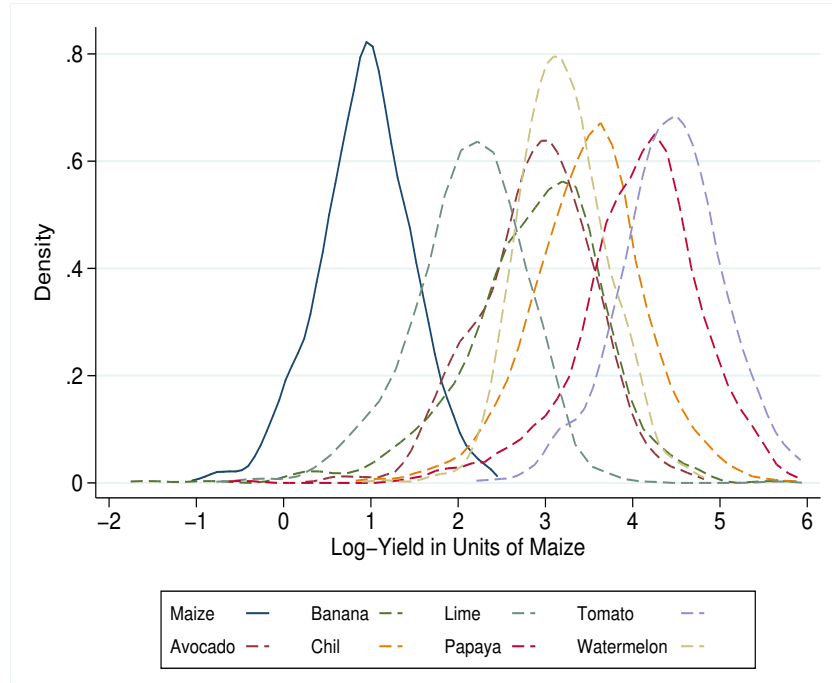
Source: Author's calculations using Input-Output Data from Mexico, 2008.

C Agricultural productivity with aggregate data

This section complements the results from Section 2.1 using aggregate data of Mexican agriculture. I compute value added per worker for different industries using Input-Output data. First, Table 12 presents labor productivity relative to non-agriculture for three categories of crops. Value added per worker in maize and other grains is less than half of value added per worker in cash crops production. Moreover, the productivity gap between agriculture and non-agriculture is almost the same as the one between maize and non-agriculture. The latter implies that the large agricultural productivity gap is actually measuring large productivity differences with respect to unproductive staple crops that have the largest employment share.

Second, to explore if the year of the ENA 2014 surveys was important for the empirical results, I compute log-yields of crops using state-panel data of agricultural production and prices in Mexico from 1980 to 2012. For each state in every year, yields are expressed in units of maize using relative prices. These yields are detrended using a linear regression with respect to time, taking 2012 as the base year. Figure 10 shows the non-parametric densities of crop yields. According to this data, the average log-yield of each of these fruits is higher than the average yield of maize; in some cases, like with tomatoes, the difference in average log-yield is around 4. These results imply that the yield distribution of many fruits first-order stochastically dominates the yield distribution of maize. Thus, these data suggests that 2014

Figure 10: Crop yields distributions



Notes: Kernel densities (Epanechnikov). Log-yields are detrended using a linear regression and taking 2012 as the base year.
Source: Author's estimates using data from SIAP. State panel from 1980 to 2012.

was not a special year and the productivity gaps between types of crops are persistent over time. If any, this motivates future research to explore other possible barriers.

D Subsistence production of crops

In this section I provide evidence on subsistence consumption for different crops. Staple crops like maize and rice are the main food component of a population's diet, especially in poor countries. A key distinction between staple crops and cash crops is that a relatively high share of staple crops production is used for subsistence requirements of food, while most cash crops production is sold to richer regions, within and outside of a country.

To analyze subsistence requirements by crop, Table 13 shows the average share of farm production used for family consumption. I distinguish between farms of all sizes and farms with less than 20 hectares (these are the ones that are not considered commercial farms). The results show that the share of maize production used for subsistence is significantly higher than any other fruit. That is, while more than one third of maize production is use for

Table 13: Farm production (percentage) used for family consumption by crop

Farms	Maize	Banana	Watermelon	Mango	Avocado	Chili	Lime	Papaya	Cucumber	Tomato
All sizes	30.2	11.7	6.0	5.7	5.4	5.3	4.8	3.6	3.0	2.0
Less than 20 ha.	36.1	13.1	9.6	6.5	6.0	9.1	5.4	4.8	4.1	2.8

Notes: Average share of farm production by crop.

Source: Author's estimates using data from INEGI: NSA 2014, Mexico.

family consumption on average, the range for other cash crops is between 2 and 12 percent of production. In fact, for most of these fruits, less than 6 percent of the production is used for subsistence. This pattern is the same for both groups of farms, so it is not exclusive to small-scale farming.