Journal of Development Economics 155 (2022) 102787

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

Land market distortions and aggregate agricultural productivity: Evidence from Guatemala☆

Braulio Britos [a], Manuel A. Hernandez [b,*], Miguel Robles [c], Danilo R. Trupkin [d]

[a] University of Minnesota, USA
[b] International Food Policy Research Institute, USA
[c] Universidad del Pacifico, Peru
[d] Universidad de Buenos Aires and IIEP (CONICET-UBA), Argentina

A R T I C L E I N F O

JEL classification:
O13
Q15
O40

Keywords:
Land market distortions
Output inefficiency
Agricultural productivity
Guatemala

A B S T R A C T

Farm size and land allocation are important factors in explaining lagging agricultural productivity in developing countries. This paper examines the effect of land market imperfections on land allocation across farmers and aggregate agricultural productivity. We develop a theoretical framework to model the optimal size distribution of farms and assess to what extent market imperfections can explain non-optimal land allocation and output inefficiency. We measure these distortions for the case of Guatemala using agricultural census microdata. We find that due to land market imperfections aggregate output is 19% below its efficient level for both maize and beans and 31% below for coffee, which are three major crops produced nationwide. We also observe that areas with higher distortions show higher land price dispersion and less active rental markets. The degree of land market distortions across areas co-variate to some extent with road accessibility, ethnicity, and education.

1. Introduction

The major role of agriculture in explaining large aggregate productivity disparities between developing and developed countries is well established (Caselli, 2005; Restuccia et al., 2008; Lagakos and Waugh, 2013). Poor countries employ most of their workers in agriculture and are much more unproductive than rich countries. As noted by Adamopolous and Restuccia (2014), farm size and land allocation are important factors in explaining this lagging agricultural productivity in poor countries. In particular, there are important differences in the size distribution of farms between rich and poor countries, where farms in poor countries have a much smaller operational scale and large farms have a significantly higher labor productivity than smaller ones. Further understanding farm size patterns, land productivity and allocation, and the drivers of these processes is critical to reduce the agricultural productivity gap in developing countries.

The objective of this study is twofold. First, we formally assess the impact of land market imperfections on the allocation of land across heterogeneous farmers and on aggregate agricultural productivity, holding other factors constant. We develop a model with an endogenous distribution of land size to characterize agricultural land (mis)allocation. Second, we quantify the magnitude of these distortions on output efficiency using the case of Guatemala as an example. We also examine potential factors associated with these distortions in the country by exploiting efficiency differences across areas, which can help to elaborate policies to improve efficiency in land markets.

The main contribution of our theoretical model is its simplicity in delivering an optimal size of farms based on a distribution of farmers in a given area with heterogeneous skills and facing varying transaction costs to operate land, and thereby quantify aggregate output inefficiencies due to sub-optimal land allocation. Our setup also permits to show the connection between transaction costs in the land market and equilibrium land price dispersion, and is tailored to generate specific model implications that can be empirically assessed based on available

☆ We thank the valuable comments of seminar participants at the European Economic Association Annual Meeting, Asian Meetings of the Econometric Society, Latin American and Caribbean Economic Association Annual Meeting, International Association of Agricultural Economists Triennial Congress, World Bank Conference on Land and Poverty, IFPRI and UNSAM. We also thank the helpful comments from Francisco Ceballos, Wenli Li, Francisco Olivet, Heitor Pellegrina, Maximo Torero, and Rob Vos. We thank the Instituto Nacional de Estadistica (INE) and the Secretaria de Seguridad Alimentaria y Nutricional (SESAN) of Guatemala for their help in accessing and collecting part of the data used in the analysis. Lastly, we would like to thank Andrew Foster and two anonymous referees for their many useful comments. All errors are our own.

* Correspondence to: 1201 Eye Street NW, Washington, DC 20005, USA.
E-mail addresses: brito032@umn.edu (B. Britos), m.a.hernandez@cgiar.org (M.A. Hernandez), lm.roblesf@up.edu.pe (M. Robles), danilo.trupkin@fce.uba.ar (D.R. Trupkin).

https://doi.org/10.1016/j.jdeveco.2021.102787
Received 20 June 2020; Received in revised form 9 November 2021; Accepted 12 November 2021
Available online 30 November 2021

data (such as the relationships between output efficiency, land price dispersion, and rental markets).

The distribution of land in Guatemala is highly skewed where 3.2% of the largest farms (over 22 hectares) control almost two thirds of the total agricultural land (Durr, 2016), and the market is segmented between a submarket for large landholdings' transactions and another for small landholdings. As noted by Lastarria-Cornhiel (2003), sale information of large holdings is passed by word-of-mouth among a limited group of people (with the resources to buy large properties), while the subdivision of large farms into smaller parcels for sale is not a regular practice; smallholders, in turn, sell or transfer parcels among themselves, frequently involving neighboring farmers or relatives that has resulted in further smaller lands over time. The limited performance of land market institutions, particularly the Registro General de la Propiedad (RGP), Registro de Informacion Catastral (RIC), and Fondo de Tierras (Fontierras), has also prevented the adequate functioning of land markets and put at risk property rights held by landholders. Land formal transactions are complex and expensive and excessively bureaucratic procedures decrease potential buyers' ability to purchase land, especially among smallholders. Several transactions of small farms, including inherited land, are completed without a formal registration (title update) nor a cadastre and this informality adds to land price and rent distortions (Carrera, 1999). In addition, multiple public and private land programs and regularization initiatives over the past decades have marginally contributed to the development of a dynamic land market due to persisting supply, demand, and institutional constraints (Gould, 2006; Gauster and Isakson, 2007). Overall, high land concentration and market segmentation, complex and costly land transfer procedures, and restricted market information and informality, represent major bottlenecks for an efficient and transparent functioning of land markets in Guatemala that are relevant to quantify.

We focus on white maize, black beans, and coffee for the analysis, which are three major crops produced nationwide and make up a large share of agricultural employment in Guatemala. The estimation results show that due to land market imperfections aggregate output is, on average, 81% of the efficient output for both maize and beans and 69% for coffee. Several robustness checks support these findings, suggesting that land market distortions play a larger (negative) role among high-value cash crops relative to staple crops. We also observe that areas with higher distortions (inefficiencies) exhibit a higher dispersion in land prices and less dynamic rental markets. Similarly, land market distortions across areas co-variate to some degree with accessibility (road connectivity), cultural aspects (ethnicity), and the level of education in the area.

The study ties into the general literature on factor misallocation across heterogeneous production units and productivity. Hopenhayn and Rogerson (1993) extend an industry equilibrium model developed in Hopenhayn (1992) and show that dismissal taxes can distort labor allocation across firms and have important welfare losses through a decrease in average labor productivity of over 2%. Restuccia and Rogerson (2008) focus on capital misallocation and show that policies creating distortions on prices faced by producers can lead to large distortions on total factor productivity (TFP) and aggregate output in the range of 30%–50%. Across countries, Gollin et al. (2014) find evidence of labor misallocation between agricultural and non-agricultural sectors. The authors use microdata for 151 countries and show that output per worker in the agricultural sector is roughly half of the value in the non-agriculture sector, and the differences are more pronounced among developing countries.[1]

Closer to our study, Restuccia and Santaeulalia-Llopis (2017) study misallocation across household farms in Malawi. The authors estimate farms' TFP using detailed household-level survey data and find that input allocation is relatively constant across farms despite large differences in TFP. Their results indicate that agricultural productivity would increase by a factor of 3.6 if inputs were reallocated efficiently. Factor misallocation in their study is linked to restricted land markets as most of the land is directly assigned by village chiefs (i.e. are not marketed); as a result, the potential gains from reallocation are 2.6 times larger for farms with no marketed land relative to farms with marketed land. However, more recently, Gollin and Udry (2021) cast doubt on the role of land misallocation on agricultural productivity using survey panel data from farms in Tanzania and Uganda. The authors develop a framework that distinguishes between measurement error, unobserved heterogeneity and potential misallocation and find that measurement error and heterogeneity together account for nearly 90% of the dispersion in measured productivity. These findings suggest that the potential efficiency gains through land reallocation across farmers may be lower than previous findings. While we implement a different approach using (less detailed) nationwide census microdata, our results for Guatemala are highly robust across regions and closer to the estimates of Gollin and Udry (2021), particularly for maize and beans.

Our paper also contributes to the literature examining potential factors correlated with misallocation. On this matter, Restuccia and Rogerson (2017) review the literature on the effect of misallocation on productivity and conclude that there is no dominant source of misallocation as multiple factors seem to contribute to the total effect (e.g., taxes and regulations, preferential market access, subsidies and market imperfections such as market power and frictions). Adamopoulos and Restuccia (2017) evaluate the role of land quality and geography on agricultural productivity differences, and find that the rich-poor agricultural yield gap is not due to land quality differences but to a lower efficiency in crop production. Chen (2017) models the effect of untitled lands, which creates misallocation, on agricultural productivity and finds that land titling can increase productivity across countries by up to 82.5%, where about half of the increase results directly from eliminating land misallocation. Chen et al. (2021) assess the role of land markets on factor misallocation in Ethiopia, where the state owns the land, and show that land rentals significantly reduce misallocation and increase agricultural productivity. Similarly, Chamberlain and Ricker-Gilbert (2016) find evidence that rental markets contribute to efficiency gains in Malawi and Zambia by facilitating the transfer of land from less-able to more-able smallholder farmers.[2]

The remainder of the paper is structured as follows. Section 2 presents the theoretical model and its implications. Section 3 describes the data and the land distribution and productivity among the selected crops in Guatemala. Section 4 quantifies and discusses the distortions and resulting inefficiencies. Section 5 concludes and provides some policy recommendations.

2. Model

We develop a model featuring an endogenous distribution of land size to characterize agricultural land (mis)allocation. Consider an economy (geographic area) populated by heterogeneous farmers, each one associated with a different location (we can interpret it as a "village") and an idiosyncratic level of productivity, all producing in a large "plateau" in the area. In each period, an agricultural good is produced in the plateau with both land (adjusted by quality) and labor, while other factors are held constant across farmers. Land misallocation

[1] Other studies that analyze the link between factor misallocation, aggregate productivity and output include Hsieh and Klenow (2009), Bartelsman et al. (2013), David et al. (2016), Bento and Restuccia (2017).

[2] The list of papers is certainly more extensive, including the broader literature on the link between land tenure, institutions, and agricultural productivity. For some recent studies, see Goldstein and Udry (2008), Besley et al. (2012), De Janvry et al. (2015), Jayne et al. (2016), Foster and Rosenzweig (2017), Henderson and Isaac (2017).

occurs because farmers face a transaction cost in the land rental market, which increases with the operated amount of land and with the distance (or lack of accessibility) to the plateau. For the case of Guatemala, the market clearing condition (at the plateau) may be delimited at the municipality or department level, and we opt for the latter as our baseline specification given that there are several cases in the country in which farmers share and/or operate land across municipalities (especially in mountain regions such as the Western Highlands and Dry Corridor).[3]

We abstract in the theoretical framework from other production factors for two main reasons. First, we are interested in characterizing the effect of land market imperfections on land allocation and output efficiency. In the spirit of Lucas (1978), if we explicitly include other factors in our modeled setup (such as machinery and equipment), the resulting distortions would naturally change due to the *span of control* (of adding more factors)[4]; yet, the magnitude of these changes ultimately depend on the assumed income shares of these other factors, which are typically small in developing countries compared to land and labor (see, e.g., Chen et al., 2021). Second, the data used in the empirical analysis do not include quantitative measures regarding capital or other potential endogenous factors (only indicator variables), which prevents us from taking advantage of modeling these production inputs. We still control though for these other factors (indicators) in the empirical section when deriving our measure of farmers' productivity and assess the sensitivity of our results to different income shares of land (i.e., importance of land relative to other factors in the production technology).

2.1. Set up

The agricultural good is produced by a farmer endowed with managerial skills s. In particular, we assume that a farmer i has the following simplified production function,

$$\tilde{y}_i = s_i(\beta_i l_i)^\alpha$$

where $\tilde{y}_i$ is the agricultural output and $(\beta_i l_i)$ represents the quality-adjusted land input, where β_i and l_i measure land quality and land size, respectively, all normalized by the amount of labor employed at the farm level. The technology is characterized by decreasing returns to scale on the ratio of (quality-adjusted) land per unit of labor. Efficiency in our setting thereby involves a set of reallocations of land per unit of labor across farms, motivated by the fact that agricultural labor is mostly supplied within the family, especially in developing countries.

Following Chen et al. (2021), we assume that land size and land quality are perfect substitutes for land input and net out the effect of land quality on output.[5] We thus define our relevant output y_i as output net of land quality,

$$y_i = \frac{\tilde{y}_i}{\beta_i^\alpha} = s_i l_i^\alpha$$

where parameter $\alpha \in (0, 1)$ captures land elasticity.

The farmer's managerial ability s follows a known time-invariant distribution with cumulative distribution function $F(s)$ and probability density function $f(s)$ with support $S = [\underline{s}, \overline{s}]$.[6] We consider a discrete number of N farmers with different managerial skills denoted by s_i and different locations (villages). This implies that i hereafter identifies both the farmer and location. We further assume that there is a large plateau in a given geographic area (department) where all farming takes place and every farmer from a different location in the area competes for land and produces at the plateau. We could additionally introduce many (heterogeneous) farmers competing for land and producing at their own location, but it is not essential for our main arguments.

Finally, we assume that farmers face a transaction cost on top of the rental price for land.[7] Let $\tau_i(l_i)$ be the transaction cost that farmer i has to pay to operate land size l_i. We assume that $\tau_i(\cdot)$ increases with land size, i.e., $\tau_i'(\cdot) > 0$, and more generally τ_i might vary with distance or easiness to access land in the plateau.[8]

These transaction costs can be justified in several ways. These costs may capture operational (transit) costs to produce in a common plateau that increase with land size (scale). They can also be interpreted as difficulties faced by farmers who demand plots that are located further away where information is scarcer, and the lack of information increases with distance or lower accessibility; the existence of asymmetries in the form of certain (market) power from insiders; transportation costs for implementing effective managerial control; among other factors. These transaction costs result in misallocations in the land market, which can be quantified in terms of welfare losses (output inefficiencies) in a given area. For simplicity, in our theoretical framework we abstract from other transaction costs that might be more difficult to eliminate in practice, such as those more closely related to geography and cultural factors, but we return to this discussion in the empirical section.

2.2. Farmer's problem

A farmer with managerial ability s_i demands land in order to maximize profits, taking the rental price q as given and subject to the non-negative constraint $l_i \geq 0$.

The farmer's problem is defined as,

$$\max_{l_i} \pi(s_i) = \left\{ s_i l_i^\alpha - q l_i - \tau_i(l_i) \right\}.$$

The optimal condition for the ith farmer is given by,

$$\alpha s_i l_i^{\alpha-1} = q + \tau_i'(l_i). \tag{1}$$

Without loss of generality, assume a quadratic transaction cost $\tau_i(l_i) = \frac{\tau_i}{2} l_i^2$.[9] Then, condition (1) becomes,

$$\alpha s_i l_i^{\alpha-1} = q + \tau_i l_i. \tag{2}$$

Lastly, for every pair of farmers i and j, we have the following optimal relative allocation of land,

$$\frac{l_i}{l_j} = \left(\frac{s_i/q_i}{s_j/q_j} \right)^{\frac{1}{1-\alpha}} \tag{3}$$

where $q_i = q + \tau_i l_i$ denotes the total renting cost of land for farmer i and $\frac{1}{1-\alpha} > 1$.

[3] The results, however, are marginally different when working at the municipality level as discussed below.

[4] As we add more factors in the production function, returns to scale would change, and with it the equilibrium land-size allocation and the aggregate measure of distortions.

[5] In the empirical analysis, we assume that all farmers in a municipality operate the same quality of land as lack of data prevents us from measuring land quality at the plot or farmer level. In Section 3.2 below we argue that this does not seem to be an unreasonable assumption.

[6] As is standard in the literature, in the model farmers know their own managerial skills.

[7] In our model farmers operate land no matter whether they own or rent the land. The emphasis of our paper is on the productivity implications of the operational-scale distribution of farms and as a consequence we abstract from the implications of ownership distribution (see., e.g., Adamopolous and Restuccia, 2014). While addressing land ownership could be relevant in other contexts, it is not central to our analysis.

[8] Note that τ_i can also be interpreted, on the margin, as a tax rate.

[9] There may be fixed costs involved as well, but we abstract from them to simplify the analysis as our central arguments do not change.

2.3. Market equilibrium

To solve for market equilibrium, we proceed as follows. The market clearing condition for the aggregate amount of land L in an area is given by,

$$L = \sum_{i=1}^{N} l_i. \tag{4}$$

Using conditions (3) and (4), we get the following expression for the individual land allocation of equilibrium,

$$l_i = \frac{(s_i/q_i)^{\frac{1}{1-\alpha}}}{\widetilde{S}} L \tag{5}$$

where $\widetilde{S} \equiv \sum_{i=1}^{N} \left(s_i/q_i\right)^{\frac{1}{1-\alpha}}$.

Then, individual output y_i becomes,

$$y_i = \frac{\left(s_i/q_i^{\alpha}\right)^{\frac{1}{1-\alpha}}}{\widetilde{S}^{\alpha}} L^{\alpha}.$$

Finally, the aggregate output results in,

$$Y = \frac{\widehat{S}}{\widetilde{S}^{\alpha}} L^{\alpha} \tag{6}$$

where $\widehat{S} \equiv \sum_{i=1}^{N} (s_i/q_i^{\alpha})^{\frac{1}{1-\alpha}}$.

2.4. The efficient allocation

Expression (6) shows the aggregate output resulting from potential inefficiencies arising from the (mis)allocation of land (what we call *actual output*). This aggregate output can be compared with a (theoretical) aggregate output that would result from a social planner who solves a simple land-allocation problem given the overall distribution of farmers' productivity (what we call *efficient output*).

In the context of our framework, the efficient aggregate output from a social-planner allocation is equivalent to the output that results from a market equilibrium without distortions. Consider, for instance, the special case in which there are no transaction costs; i.e., $\tau_i = 0$ for all i. In this case, the total rental cost of land becomes q for all farmers as land-market imperfections disappear.

Formally, we can define farmer i's efficient land allocation as l_i^*. Then, provided that $q_i = q$ for all i, we obtain from (5) the following expression for the efficient individual land size,

$$l_i^* = \frac{s_i^{\frac{1}{1-\alpha}}}{S} L \tag{7}$$

where $S \equiv \sum_{i=1}^{N} s_i^{\frac{1}{1-\alpha}}$. The equilibrium price q^* is further given by,

$$q^* = \alpha \left(\frac{S}{L}\right)^{1-\alpha}. \tag{8}$$

From expression (7), each farmer's efficient output is given by,

$$y_i^* = \frac{s_i^{\frac{1}{1-\alpha}}}{S^{\alpha}} L^{\alpha}.$$

Summing up over the distribution of farmers in a given area, we obtain an expression for the aggregate (economy-level) efficient output equal to,

$$Y^* = S^{1-\alpha} L^{\alpha}. \tag{9}$$

Finally, by comparing the actual output Y defined in (6) with the efficient output Y^* defined in (9), the ratio Y/Y^* offers a measure of the degree of output inefficiencies in an area resulting from land misallocation.

Using (6) and (9) we can explicitly derive the following output *efficiency ratio*,

$$\frac{Y}{Y^*} = \frac{\widehat{S}\widetilde{S}^{-\alpha}}{S^{1-\alpha}} = \frac{\left[\sum_{i=1}^{N}\left(\frac{s_i}{q_i^{\alpha}}\right)^{\frac{1}{1-\alpha}}\right]\left[\sum_{i=1}^{N}\left(\frac{s_i}{q_i}\right)^{\frac{1}{1-\alpha}}\right]^{-\alpha}}{\left[\sum_{i=1}^{N} s_i^{\frac{1}{1-\alpha}}\right]^{1-\alpha}}. \tag{10}$$

If $\tau_i = 0$ for all i (i.e. no land market imperfections), it can easily be shown that $\widehat{S} = \widetilde{S} = S$ and $Y = Y^*$, which results in an efficiency ratio equal to 1.

We return to this discussion in Section 4 where we calculate the above ratio to quantify the magnitude of output inefficiencies for selected crops across Guatemala. We also consider an adjusted benchmark in the empirical section.

2.5. Model implications

Characterizing the equilibrium results under the distorted (actual) and efficient (theoretical) land allocation allows us to find several appealing model implications.

First, we rewrite Eq. (3); i.e., the relative (distorted) land allocation between farmers i and j, as follows,

$$\frac{l_i}{l_j} = \left(\frac{s_i}{s_j}\right)^{\frac{1}{1-\alpha}} \left(\frac{q_j}{q_i}\right)^{\frac{1}{1-\alpha}}. \tag{11}$$

Expression (11) suggests that the relative land size of farmer i increases with managerial ability (s_i/s_j) and decreases with the ratio of renting costs (q_i/q_j), which is a function of the relative distortions (i.e., the transaction costs faced by each farmer).

Likewise, the relative efficient land allocation (without distortions) between farmers i and j is given by,

$$\frac{l_i^*}{l_j^*} = \left(\frac{s_i}{s_j}\right)^{\frac{1}{1-\alpha}}. \tag{12}$$

This expression shows that the relative efficient land size between farmers i and j solely depends on managerial abilities (s_i/s_j).

Combining Eqs. (11) and (12), we can derive a relationship between the two land allocations from which we can make some inference about the size distribution of farms. Let $z_{ij} \equiv l_i/l_j$ and $z_{ij}^* \equiv l_i^*/l_j^*$. Then,

$$\frac{z_{ij}}{z_{ij}^*} = \left(\frac{q_j}{q_i}\right)^{\frac{1}{1-\alpha}}. \tag{13}$$

In general, we can characterize the following model implications from Eqs. (3), (12) and (13):

1. The efficient land size of less productive farmers will be lower than the efficient land size of more productive ones.
2. The actual (likely distorted) land size of less productive farmers will not necessarily be lower than the actual land size of more productive ones.
3. The higher the transaction cost τ for farmer i, the lower her relative land size in equilibrium (z_{ij}) compared to the efficient ("ought-to-be") land size (z_{ij}^*).
4. The higher the land market distortions (the higher the τ's across farmers), the higher the price dispersion in the market (i.e., the higher the differences between any q_i and q_j) and the less efficient the allocation of land.

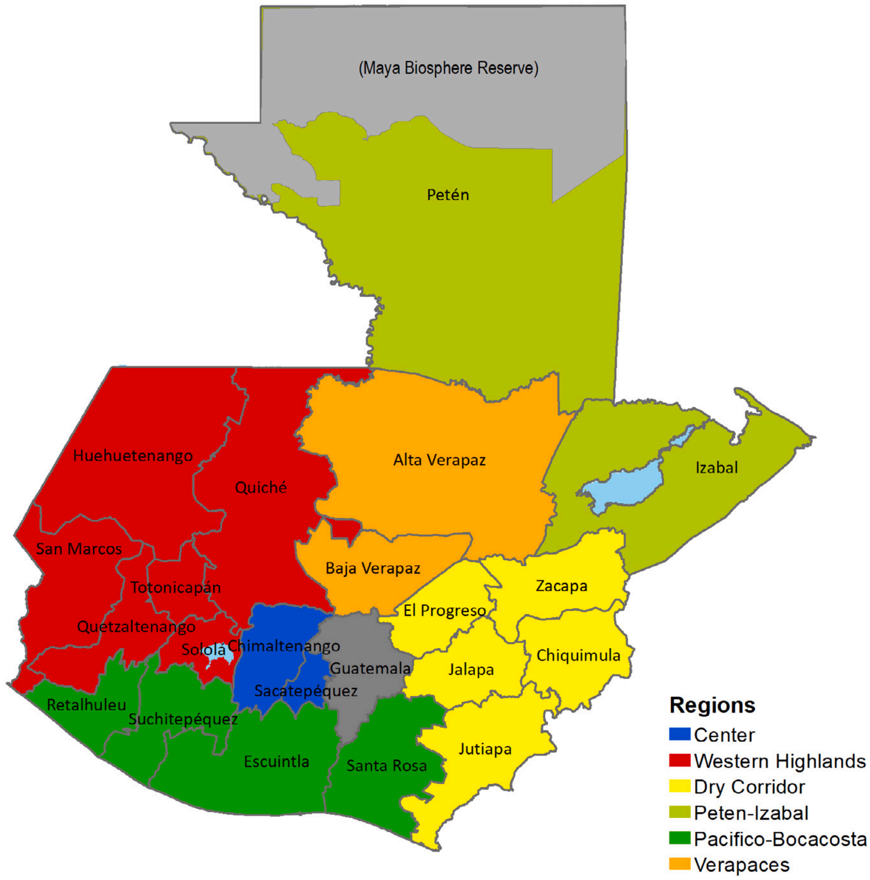

Fig. 1. Map of Guatemala and regions considered.

3. The case of Guatemala

Guatemala is an interesting case study as it exhibits a large degree of heterogeneity in terms of climate, geography, ethnic composition, and rural development. There is a wide variation of agricultural activities; from large/medium- to low-scale farming and from high-value export crops such as coffee and sugar cane to food staple crops such as maize and beans. Land markets, in turn, are characterized by a high degree of concentration and segmentation with limited market information, high informality, and complex and costly transaction procedures, as noted earlier.

For the analysis below, we group the 22 departments of the country into six major geographic regions as shown in the map in Fig. 1: Center, Western Highlands ('Altiplano Occidental'), Dry Corridor ('Corredor Seco'), Peten-Izabal, Pacífico-Bocacosta, and Verapaces. We exclude the department of Guatemala from the central region since the capital city is located in this department and there is a much lower presence of agricultural activities relative to other activities, as opposed to the other departments.[10] The departments within each region share similar socioeconomic, accessibility, and agro-climatic conditions.

3.1. Data

The dataset used in the analysis is the microdata from the last census of agriculture in Guatemala, 'IV Censo Nacional Agropecuario 2003', corresponding to the crop year 2002–03 collected by the National Statistical Institute (INE). The census includes information on land size and use (for crops, cattle farming and other activities), production, labor and input use, machinery and equipment ownership, farmers' socioeconomic characteristics and geographic location. We focus on white maize, black beans, and coffee, which results in a working sample of 396,317, 113,133, and 147,353 producers for each crop.[11]

Maize, beans, and coffee are three major crops produced nationwide in Guatemala and together generate 62% of the agricultural employment (MAGA, 2011; MAGA, 2013). Maize, specially white maize, is by far the most common and extended food crop produced in the country with an annual harvested area of 841,094 hectares (Ha) and total production of 1,672,527 metric tonnes (MT) as of 2011/12; the major producer regions are Peten-Izabal (where maize production is combined with cattle farming activities), Western Highlands, and Verapaces. Beans is the second major staple crop, which is mainly produced for self-consumption and local markets across the country, with an annual harvested area of 238,140 Ha and production of 199,946 MT; the major producer regions are the Dry Corridor, Peten-Izabal, and Western Highlands. Coffee is the second major export crop and is produced in multiple regions with an annual harvested area of 252,415 Ha and production of 245,752 MT; the major producer regions include Western Highlands, Pacífico-Bocacosta, and Verapaces.[12] Focusing on these crops allows as to make comparisons across regions as well as to assess whether inefficiencies resulting from potential land misallocation

[10] The estimation results, however, are not sensitive to including this department.

[11] Around 2.5% of producers from the raw census data are excluded from the analysis due to missing observations, likely typos, and extreme values for key variables of interest.

[12] Sugar cane is the main export crop in Guatemala but its production, which is basically large-scale farming with a total harvested area of 239,261 MT, is concentrated in a specific region (Pacífico-Bocacosta), reason why we exclude it from the present study.

Table 1
Size distribution of landholdings devoted to agriculture and by crop.

Landholding size	All regions	Center	Western Highlands	Dry Corridor	Peten-Izabal	Pacifico-Bocacosta	Verapaces
Less than 1 Ha	64.7%	83.9%	79.9%	53.2%	17.7%	60.3%	39.4%
1-2 Ha	17.2%	11.9%	12.0%	27.1%	17.5%	18.2%	26.2%
2-5 Ha	11.9%	3.3%	5.8%	15.3%	29.0%	15.2%	23.6%
5-10 Ha	2.9%	0.4%	1.4%	2.5%	11.1%	2.9%	6.2%
10-20 Ha	1.6%	0.2%	0.6%	1.0%	9.4%	1.4%	2.9%
More than 20 Ha	1.7%	0.3%	0.3%	0.8%	15.3%	2.1%	1.6%
White maize							
Less than 1 Ha	73.3%	95.5%	90.5%	70.8%	20.9%	66.3%	65.1%
1-2 Ha	15.1%	3.6%	7.0%	20.2%	25.1%	18.0%	25.4%
2-5 Ha	9.3%	0.7%	2.2%	7.8%	40.4%	12.8%	8.8%
5-10 Ha	1.6%	0.1%	0.2%	0.9%	10.2%	1.9%	0.6%
10-20 Ha	0.5%	0.0%	0.1%	0.3%	2.7%	0.7%	0.1%
More than 20 Ha	0.2%	0.0%	0.0%	0.1%	0.7%	0.4%	0.1%
Black Beans							
Less than 1 Ha	82.1%	97.4%	97.0%	79.1%	43.7%	92.3%	94.7%
1-2 Ha	10.2%	2.1%	2.3%	14.4%	25.0%	5.4%	3.9%
2-5 Ha	6.5%	0.5%	0.6%	5.5%	25.8%	1.7%	1.2%
5-10 Ha	1.0%	0.0%	0.1%	0.7%	4.2%	0.4%	0.1%
10-20 Ha	0.2%	0.0%	0.0%	0.2%	1.0%	0.1%	0.0%
More than 20 Ha	0.1%	0.0%	0.0%	0.1%	0.3%	0.1%	0.0%
Coffee							
Less than 1 Ha	81.0%	78.7%	82.4%	75.4%	74.8%	62.8%	92.6%
1-2 Ha	10.6%	13.4%	11.1%	12.1%	13.2%	16.9%	4.6%
2-5 Ha	5.8%	4.0%	4.7%	9.1%	9.4%	13.6%	1.9%
5-10 Ha	1.2%	1.3%	0.8%	2.1%	1.6%	3.0%	0.4%
10-20 Ha	0.5%	0.8%	0.3%	0.8%	0.3%	1.3%	0.2%
More than 20 Ha	0.8%	1.8%	0.7%	0.5%	0.7%	2.5%	0.3%

Note: Calculations based on size of landholdings dedicated to agriculture in the top panel and to the production of white maize, black beans and coffee in the other panels. The Center region includes the departments of Sacatepequez and Chimaltenango; Western Highlands includes Huehuetenango, Quiche, San Marcos, Quetzaltenango, Totonicapan and Solola; Dry Corridor includes Chiquimula, Jutiapa, Jalapa, El Progreso and Zacapa; Peten-Izabal includes Peten and Izabal; Pacifico-Bocacosta includes Retalhuleu, Suchitepequez, Escuintla and Santa Rosa; and Verapaces includes Alta Verapaz and Baja Verapaz.

Table 2
Average and dispersion of yields per worker by land size.

	All regions		Center		Western Highlands		Dry Corridor		Peten-Izabal		Pacifico-Bocacosta		Verapaces	
	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha
White Maize														
Mean	16.0	12.0	18.2	10.4	15.7	11.3	13.3	7.9	18.4	14.0	28.0	19.0	11.3	9.1
St. Dev.	15.6	12.4	16.6	10.5	14.4	11.9	13.6	9.0	15.4	12.5	22.5	16.5	11.4	10.2
IQR (P75-P25)	16.2	12.3	15.2	9.8	17.2	11.1	12.6	7.0	17.2	14.2	25.8	15.7	11.7	9.4
Observations	289,338	106,979	30,669	1,460	138,173	14,351	39,841	16,448	9,091	35,653	24,181	12,836	47,383	26,231
Black Bean														
Mean	3.9	3.5	4.0	2.9	4.1	3.0	4.0	2.5	4.8	4.4	5.1	2.6	3.2	2.4
St. Dev.	3.8	3.4	3.7	3.0	4.0	3.7	3.8	2.7	3.8	3.7	4.7	2.8	3.3	2.6
IQR (P75-P25)	3.9	3.6	3.0	3.0	4.3	2.8	3.7	2.3	5.2	4.4	5.2	2.8	3.0	2.4
Observations	92,650	20,483	8,121	221	21,673	629	26,774	7,078	8,397	10,868	4,468	379	23,217	1,308
Coffee														
Mean	16.0	9.6	17.3	11.3	16.7	8.2	16.5	9.7	9.3	11.3	24.5	12.5	11.1	6.4
St. Dev.	21.9	15.0	21.0	14.9	21.9	12.3	23.0	16.3	13.8	19.4	30.1	18.3	16.0	10.1
IQR (P75-P25)	15.5	8.9	17.6	10.4	18.7	7.7	16.1	8.8	10.0	10.8	22.9	11.2	10.5	5.8
Observations	119,306	28,047	5,702	1,543	53,388	11,367	16,539	5,481	599	183	11,745	6,987	31,333	2,486

Note: The Center region includes the departments of Sacatepequez and Chimaltenango; Western Highlands includes Huehuetenango, Quiche, San Marcos, Quetzaltenango, Totonicapan and Solola; Dry Corridor includes Chiquimula, Jutiapa, Jalapa, El Progreso and Zacapa; Peten-Izabal includes Peten and Izabal; Pacifico-Bocacosta includes Retalhuleu, Suchitepequez, Escuintla and Santa Rosa; and Verapaces includes Alta Verapaz and Baja Verapaz. The Interquartile Range (IQR) is equal to the difference between the 75-th percentile and 25-th percentile values. Yields are expressed in quintals per hectare (per worker).

are more acute for certain types of crops, i.e. crops that involve small versus small/medium scale production, staple crops for subsistence and local markets versus cash crops for external markets.

Table 1 provides general descriptive statistics of the land size distribution in Guatemala. The top panel of the table shows the size distribution of landholdings dedicated to agricultural activities in the whole country and disaggregated by region. Farms under one hectare, considered as "infra-subsistence" farms by the Ministry of Agriculture (INE, 2006), comprise close to 65% of the landholdings in Guatemala whereas very large farms (over 20 hectares) comprise less than 2% of the landholdings.[13] The small size of landholdings is a regular pattern in

[13] As noted by Durr (2016), while the vast majority of farms in Guatemala are small, close to two thirds of the agricultural land in the country correspond to large-scale farms.

Table 3
Average and dispersion of productivity per worker by land size.

	All regions		Center		Western Highlands		Dry Corridor		Peten-Izabal		Pacifico-Bocacosta		Verapaces	
	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha	< 1 Ha	>= 1 Ha
White Maize														
Mean	−0.1	0.3	0.0	0.8	0.0	0.5	−0.1	0.4	−0.5	0.1	−0.2	0.4	−0.3	0.3
St. Dev.	0.7	0.7	0.6	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.7	0.7	0.6	0.7
IQR (P75-P25)	0.9	0.9	0.8	0.9	0.9	0.9	0.8	0.8	0.8	0.8	0.9	0.8	0.9	0.9
Observations	289,338	106,979	30,669	1,460	138,173	14,351	39,841	16,448	9,091	35,653	24,181	12,836	47,383	26,231
Black Bean														
Mean	−0.2	0.4	0.0	1.0	−0.1	1.1	−0.2	0.4	−0.6	0.3	−0.3	0.5	−0.2	0.9
St. Dev.	0.8	0.7	0.7	0.7	0.8	0.8	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.9
IQR (P75-P25)	1.0	1.0	0.9	1.0	1.0	1.1	0.9	0.9	0.9	0.9	1.0	0.9	1.1	1.1
Observations	92,650	20,483	8,121	221	21,673	629	26,774	7,078	8,397	10,868	4,468	379	23,217	1,308
Coffee														
Mean	0.0	0.8	0.1	0.8	0.0	0.8	−0.1	0.9	0.1	1.0	−0.1	0.8	−0.1	1.1
St. Dev.	1.0	1.1	1.0	1.0	0.9	1.0	1.1	1.1	1.3	1.6	1.0	1.0	1.0	1.2
IQR (P75-P25)	1.3	1.3	1.3	1.2	1.2	1.2	1.5	1.4	1.8	2.7	1.3	1.3	1.3	1.5
Observations	119,306	28,047	5,702	1,543	53,388	11,367	16,539	5,481	599	183	11,745	6,987	31,333	2,486

Note: Farmers productivity expressed in natural logarithms, derived from the full-sample estimations depicted in Table A.2 in the Appendix. The Center region includes the departments of Sacatepequez and Chimaltenango; Western Highlands includes Huehuetenango, Quiche, San Marcos, Quetzaltenango, Totonicapan and Solola; Dry Corridor includes Chiquimula, Jutiapa, Jalapa, El Progreso and Zacapa; Peten-Izabal includes Peten and Izabal; Pacifico-Bocacosta includes Retalhuleu, Suchitepequez, Escuintla and Santa Rosa; and Verapaces includes Alta Verapaz and Baja Verapaz. The Interquartile Range (IQR) is equal to the difference between the 75-th percentile and 25-th percentile values.

developing countries as opposed to developed countries, where a large share of farms operate under much larger scales.[14]

The table further shows a large variation in the size distribution of landholdings across regions. For instance, 84% and 80% of the agricultural landholdings in the Central region and Western Highlands are smaller than one hectare; the case of the former is explained by the lower agricultural development in the central part of the country, while in the case of the latter this region is the poorest in terms of economic and rural development and there is a large presence of smallholder, subsistence agriculture. In contrast, the Petén-Izabal region combines agricultural with cattle farming activities and thus exhibits larger landholdings than the rest of the country.

The lower panels of the table present the land size distribution for the crops of interest. We generally observe a larger prevalence of smaller-scale farming in beans production across regions, relative to maize and coffee. The average landholding size dedicated to beans is 0.71 hectares versus 1.02 hectares for maize and 1.47 hectares for coffee.

Table 2 presents summary statistics of yields per worker by land size (for less than and more than one hectare) for each crop and region. Yields are a standard measure of agricultural land productivity defined as production (in quintals) per hectare.[15] Two interesting patterns emerge from the table. First, small farms mostly exhibit higher (and more dispersed) yields than large farms, which is indicative of decreasing returns to scale. This inverse relationship between land size and yields is commonly found in the literature (Barrett, 1996; Place, 2009; Barrett et al., 2010). Second, there are large differences in both the average and dispersion of yields across regions by crop, where Pacífico-Bocacosta seems to be the region with the highest yields and Verapaces the region with the lowest yields.

[14] See Lowder et al. (2016, 2019) for an extensive comparison of farm size and distribution across low-, lower-middle, upper-middle and high-income countries. Restuccia and Santaeulalia-Llopis (2017) report, for instance, that more than 81% and 46% of farms in the United States (US) and Belgium have more than 10 hectares, and only 15% of farms in Belgium have less than one hectare and none in the US.

[15] We divide yields per worker for comparison purposes with our productivity measure derived below, which is normalized per worker following our theoretical setup (as we are interested on assessing reallocation of land per unit of labor across farmers). One quintal is equivalent to 100 pounds or 46 kilograms.

In Fig. 2, we report scatterplots of average landholding size against total (per capita) volume of production across departments by crop. We observe that departments with larger production volumes are generally those that, on average, exhibit larger landholdings. This apparent positive relationship between production and land size is similarly depicted in Fig. A.1 in the Appendix that plots the share of small landholdings (under one hectare) and very large landholdings (above 20 hectares) in each department against their per capita volume of production: we find that in departments with a higher share of "infra-subsistence" farms (less than one hectare) their production volume across three crops seems to be smaller than among departments with a lower share of small farms, while in departments with a higher share of farms over 20 hectares their production volume (except for beans) is somewhat larger than among departments with a smaller share of very large farms. These aggregate production and land size patterns at the department level are overall in line with the cross-country evidence presented in Adamopolous and Restuccia (2014) and indicative of potential land market distortions (inefficiencies).

3.2. A measure of farmer productivity

We now turn to the calculation of our productivity measure at the farmer level for white maize, black beans and coffee using a regression-based approach. Following the theoretical setup, we are interested in deriving farmer-level productivity (ability) measures that permit to compare hypothetical (efficient) land allocations with actual allocations to quantify the resulting output inefficiencies.

We implement a two-stage approach. First, following Restuccia and Santaeulalia-Llopis (2017), Chen et al. (2021), we derive a measure of managerial ability s for each farmer i producing each crop based on the production function defined in Section 2, reproduced here for convenience,

$$y_i = s_i l_i^{\alpha}. \tag{14}$$

Using expression (14), we solve for s_i and take the natural logarithm of both sides to get the first-stage residual,

$$\ln s_i = \ln y_i - \alpha \ln l_i. \tag{15}$$

We use the census microdata on output and land size per unit of labor. We consider three different values for α (0.2, 0.3, and 0.4) based on the range of land income shares estimated by Valentinyi and

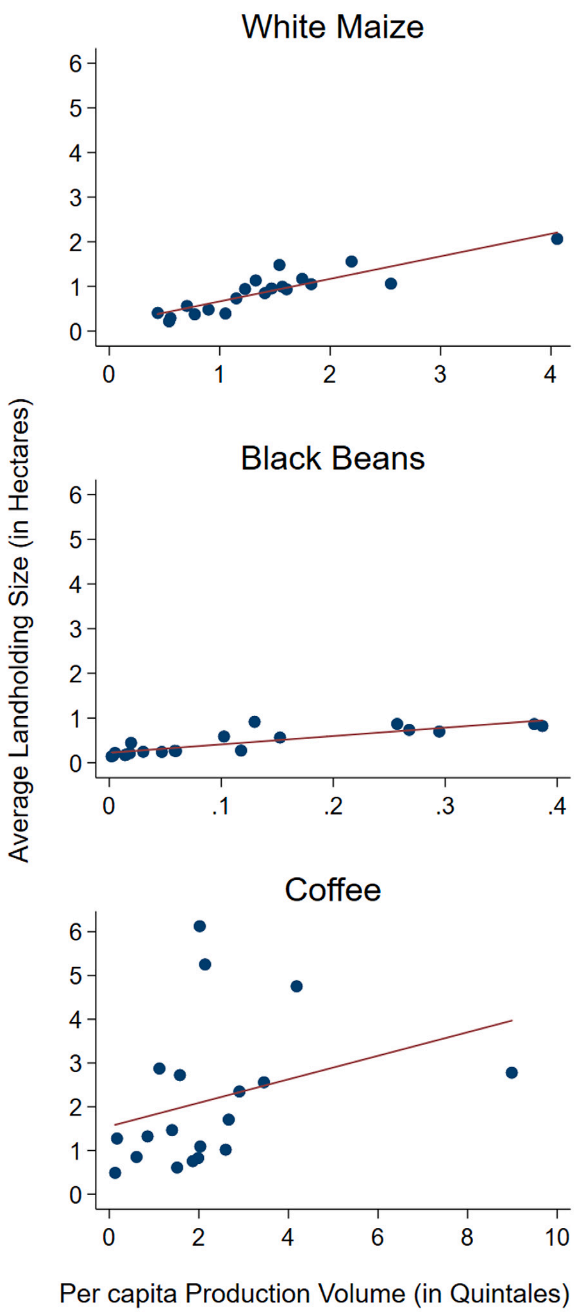

Fig. 2. Average landholding size and per capita production volume across departments by crop. Note: Landholding size is defined as the number of hectares dedicated to the production of each crop. The volume of production is measured in Guatemalan quintales where one quintal is equivalent to 100 pounds.

Herrendorf (2008) for the US ($\alpha = 0.18$) and Restuccia and Santaeulalia-Llopis (2017) for Malawi ($\alpha = 0.39$); the larger α value in Malawi is explained by the lower level of mechanization in the agriculture sector relative to the US. The mid-point value of $\alpha = 0.3$ could be regarded as our benchmark (reference) value given the general level of agricultural development in Guatemala, which may still vary by crop.[16]

Following our theoretical set up, the output measure y_i should also be adjusted by land quality; i.e., $y_i = \tilde{y}_i/\beta_i^{\alpha}$. We accordingly assume that β_i is the same across all farmers within their municipality as lack of data prevents us from measuring land quality at the farmer level.[17] We find support for this assumption using a secondary household dataset that covered more than half of the municipalities in the country (described in detail in Section 4.3). As shown in Panel A of Table A.1 in the Appendix, self-reported land quality dispersion is relatively low at the municipality level: the standard deviation of an index that measures self-reported farmers' land-quality compared to the "best" land-quality in their municipality is, on average, 1.64 in a 1-to −10 scale across municipalities, while the coefficient of variation is 0.26; likewise, Panel B of the table shows that the variance of self-reported land prices (that can serve as a proxy of land quality) among farmers within a municipality is lower than the variance of land prices among farmers of municipalities that belong to the same department or region.

Second, we regress the derived s_i series from the first step on a set of observable factors not included in the modeled setup to further control for likely heterogeneity in the use of these other factors that could affect our measure of productivity. These controls include farmer's years of education, share of family labor force, use of machinery and equipment, if farmer has livestock, use of enhanced seeds, fertilizer, and pesticides, if farm has an irrigation system, and number of different crops cultivated as a proxy of specialization. We also include farmer's age and gender to control for the eventual ascendancy and advantage position of older and male farmers in rural Guatemala, particularly on Mayan cultures, which could be correlated with (asymmetric) input access and land allocation patterns beyond a farmer's inherent ability; we then put back these two exogenous variables into our productivity measure. In addition, we include fixed effects by municipality to control for (unobserved) differences across municipalities, including biophysical, accessibility, cultural, and social factors that could contribute to productivity differences, as well as for land quality that is assumed to be similar within each municipality.[18] It is also worth noting that the lack of detailed georeferenced data with the exact location of farmers prevents us from controlling for production (weather) shocks at the farmer level, but we are not aware of any extreme event in Guatemala that could have significantly affected agricultural activities during our period of analysis; both in 2003 and in the previous year there were no major droughts, floods, or frosts across the country (see, e.g., Bardales et al., 2019; BID, 2019).

We report in Table A.2 in the Appendix the regression results of this second step for the full sample of farmers by crop.[19] The coefficients of the control variables generally have the expected signs. We find a positive correlation across all three crops between our s measure and

[16] We approximated a value for Guatemala using the 2014 National Survey of Living Conditions (ENCOVI), which is the best available secondary data source that has information, although still very limited, on agricultural output and sales, land rentals, and certain intermediate input costs (seeds, fertilizers, pesticides). We find a mean ratio of implied land costs to farm output value of 0.29-0.30 based on two different calibration methods, which provides support to our mid-point value of 0.3. Further details are available upon request.

[17] We have access to a land quality index at the municipality level, reported by the Ministry of Agriculture, equal to the share of high-quality land for agricultural activities (based on a set biophysical conditions), but this assumption is, in practice, embedded in the inclusion of municipality fixed effects in the second-stage regression discussed below.

[18] We considered an alternative specification where we interacted a land quality index, available at the municipality level, with all control variables to account for eventual heterogeneity in input, technology, and land decisions by land quality, but the results are qualitatively similar to our base findings.

[19] In the analysis below we perform separate regressions by department in order to further account for potential heterogeneity across departments. All regressions include an indicator variable that takes the value of one when the farmer's gender, age, and years of education are missing (as these variables were not always reported), and zero otherwise. For these cases, we assign the median value at the municipality level. The estimated distortions are though not sensitive to not imputing missing values for these three variables (which results in fewer observations for the analysis) or excluding these variables from the regressions. Further details are available upon request.

farmers' education, use of equipment, presence of irrigation system, and higher degree of specialization (i.e., smaller number of different crops produced). We similarly observe positive correlations between s and use of machinery, fertilizer, and pesticide for some crops. A larger share of family labor force is positively associated with s for maize and beans production. Female producers, in contrast, are associated with a lower productivity as well as younger farmers (except for beans).

The measure of farmer productivity used for the analysis is the residual of this second-stage regression augmented by age and gender (interacted with their corresponding estimated coefficients). Fig. A.2 in the Appendix plots the corresponding regional distributions (kernel densities) of the estimated productivities by crop. It is clear that the distributions are very similar across regions, which is indicative of alike farmer-level heterogeneities across locations.[20] These comparable distribution patterns permit to infer that potential differences in output efficiencies across regions, discussed in the next section, are not necessarily driven by heterogeneity differences across locations.

Table 3 reports summary statistics of the derived farmer productivities (per worker) by land size for each crop and region. In contrast to the yields presented in Table 2, farmer productivities are on average higher among farms of one or more hectares than among farms of less than one hectare. These opposite relationships between yields and productivity with land size provide additional support that our derived productivity measure is essentially capturing managerial skills and are in line with the results of Aragon et al. (2019) for Uganda; as noted by the authors, yields may capture farm productivity combined with decreasing returns to scale (and market imperfections). Overall, the positive correlation between our estimated farmer productivity measure and land size suggests that land allocation is related in some degree to a farmer's ability.

We still acknowledge though that we are constrained to cross-sectional data that prevent us from implementing alternative total factor productivity approaches that rely on panel data methods. Similarly, data limitations restrict us from modeling a richer production function (with additional factors) that could allow us to estimate our measure of farmer productivity in one step; i.e., we know from the census microdata whether farmers have machinery or use fertilizer, but we do not know the value of their equipment or amount of input use. We discuss below the potential influence of measurement error in our estimations. Next, we evaluate to what extent the current allocation of land is efficient.

4. Quantitative analysis

In this section, we quantify the magnitude of land misallocation in Guatemala for maize, beans, and coffee. We first assess how the actual land allocation for each crop compares with a benchmark, efficient allocation chosen by a hypothetical social planner based on the estimated productivities. We then consider an adjusted benchmark. We also examine whether the resulting inefficiencies are in line with some of the model implications outlined in Section 2. Finally, we evaluate potential channels that may explain the observed distortions across areas.

4.1. Output efficiency ratios

Our approach is built on Restuccia and Santaeulalia-Llopis (2017). First, we solve a simple optimization problem for a hypothetical social planner intending to maximize aggregate output by allocating land according to the distribution of farmers' estimated productivities. As discussed earlier, the solution to this problem results in the same aggregate output as the one described in expression (9) under market equilibrium without distortions. Second, we compare this efficient aggregate output with the output that results from the land size distribution found in the data, which is characterized in expression (6).

The efficient land allocation in a given department is in practice obtained by solving the following social planner problem,

$$Y^* = \max_{\{l_i\}} \sum_{i=1}^{N} s_i l_i^{\alpha}, \quad \text{subject to} \quad L = \sum_{i=1}^{N} l_i$$

where Y^* denotes the efficient output.

The solution to the optimization problem is straightforward as the marginal product of land must be equal across farmers. The following expression is equivalent to expression (7), and represents the efficient land allocation of an individual farmer,

$$l_i^* = \frac{s_i^{\frac{1}{1-\alpha}}}{S} L$$

where $S \equiv \sum_{i=1}^{N} s_i^{\frac{1}{1-\alpha}}$, as defined above.

Hence, the optimal land size of each farmer depends on her estimated productivity relative to the whole distribution of productivities. Letting $S_i \equiv \frac{s_i^{\frac{1}{1-\alpha}}}{S}$, it follows that,

$$Y^* = \sum_{i=1}^{N} s_i \left(S_i L\right)^{\alpha} = S^{1-\alpha} L^{\alpha},$$

which is equivalent to expression (9).

Lastly, we compare the efficient output with the output under the current land allocation, defined as

$$Y = \sum_{i=1}^{N} s_i l_i^{\alpha}$$

where l_i is directly observed in the census microdata.

We thus calculate efficiency ratios Y/Y^* for each department and crop where the underlying assumption is that all farmers producing a certain crop in a given department could eventually rent in/out land within their department. These ratios are constructed based on the distribution of farmers' productivities estimated at the department level. Table 4 presents the corresponding efficiency ratios Y/Y^* by crop and region where the reported ratios are the averages of the estimated departmental ratios in a region weighted by the number of producers in each department. A higher efficiency ratio indicates that the current land allocation is closer to the optimal allocation from a social planner's perspective. We provide results for α values of 0.2, 0.3, and 0.4, where a larger value of α implies a lower level of mechanization in the production process.

The table shows varying degrees of inefficiencies among the selected crops that can be attributed to land misallocation (all else equal). For $\alpha = 0.3$, our mid-point value, the efficiency ratio ranges between 78.8% and 83.5% for white maize across regions, between 79.8% and 83.5% for black beans, and between 59.8% and 71.6% for coffee. Interestingly, the regions with the largest efficiency ratios are generally the regions where a significant share of the production of each crop concentrates (except Peten-Izabal for coffee), while the region with the highest yields (Pacífico-Bocacosta) shows one of the lowest efficiency ratio.

Overall, we find a larger output efficiency for maize and beans, which are staple crops with a higher prevalence of small-scale and subsistence agriculture, relative to coffee, which is a high-value export crop characterized by small- and medium-scale production. The average efficiency ratio for both maize and beans is around 81%, which implies an aggregate output gap between the current land allocation and the theoretically efficient allocation of one fifth; the average efficiency ratio for coffee is roughly 69%, which implies an output gap of about one third. In the hypothetical case that all land market distortions were removed, total output would increase to a lower extent for white

[20] The only exception is Peten-Izabal for coffee, which is precisely the crop-region where we have fewer observations; there are only 782 coffee producers in Peten-Izabal compared to 4847–152,524 producers in the other crop-region pairs, as reported in Table 2.

Table 4
Efficiency ratios (Y/Y^*).

Region	White Maize			Black Beans			Coffee		
	α = .2	α = .3	α = .4	α = .2	α = .3	α = .4	α = .2	α = .3	α = .4
All regions	86.7%	81.1%	75.8%	86.6%	80.8%	75.5%	77.4%	68.7%	61.3%
Center	87.5%	82.0%	76.9%	87.8%	82.4%	77.4%	69.7%	59.8%	52.2%
Western Highlands	86.2%	80.5%	75.3%	86.0%	80.1%	74.6%	78.1%	69.7%	62.3%
Dry Corridor	86.8%	81.1%	75.6%	86.1%	80.2%	74.8%	78.4%	69.8%	62.2%
Peten-Izabal	88.6%	83.5%	78.7%	88.5%	83.5%	78.8%	79.9%	71.6%	63.9%
Pacifico-Bocacosta	84.8%	78.8%	73.6%	86.0%	80.1%	74.7%	74.4%	65.2%	57.8%
Verapaces	87.2%	81.5%	76.0%	85.8%	79.8%	74.2%	78.7%	70.1%	62.6%

Note: The regional ratios reported are the corresponding averages of the departmental ratios in a region weighted by the number of producers in each department. The ratio for All regions is the weighted average across regions. The Center region includes the departments of Sacatepequez and Chimaltenango; Western Highlands includes Huehuetenango, Quiche, San Marcos, Quetzaltenango, Totonicapan and Solola; Dry Corridor includes Chiquimula, Jutiapa, Jalapa, El Progreso and Zacapa; Peten-Izabal includes Peten and Izabal; Pacifico-Bocacosta includes Retalhuleu, Suchitepequez, Escuintla and Santa Rosa; and Verapaces includes Alta Verapaz and Baja Verapaz.

Table 5
Baseline and adjusted output gaps ($1 - Y/Y^*$).

Region	White Maize		Black Beans		Coffee	
	Baseline	Adjusted	Baseline	Adjusted	Baseline	Adjusted
All regions	18.9%	15.1%	19.2%	14.9%	31.3%	25.8%
Center	18.0%	13.0%	17.6%	12.9%	40.2%	31.3%
Western Highlands	19.5%	17.2%	19.9%	17.9%	30.3%	27.0%
Dry Corridor	18.9%	11.9%	19.8%	12.6%	30.2%	20.1%
Peten-Izabal	16.5%	14.2%	16.5%	14.2%	28.4%	25.1%
Pacifico-Bocacosta	21.2%	12.1%	19.9%	11.1%	34.8%	24.6%
Verapaces	18.5%	16.0%	20.2%	17.5%	29.9%	26.7%

Note: The regional output gaps reported are the corresponding averages of the departmental output gaps in a region weighted by the number of producers in each department. The first column for every crop represents the output gap from our baseline estimation (Table 4). The second column reports an adjusted gap where the departmental output gaps are computed by comparing the estimated departmental efficiencies calculated in our base results with an adjusted degree of efficiency that incorporates the difference in the share of rented agricultural land in each department relative to Escuintla, which is the department with the most dynamic land rental market in the country (see formula in footnote 24). The output gap for All regions is the weighted average across regions. The output gaps are derived assuming $\alpha = 0.3$. The Center region includes the departments of Sacatepequez and Chimaltenango; Western Highlands includes Huehuetenango, Quiche, San Marcos, Quetzaltenango, Totonicapan and Solola; Dry Corridor includes Chiquimula, Jutiapa, Jalapa, El Progreso and Zacapa; Peten-Izabal includes Peten and Izabal; Pacifico-Bocacosta includes Retalhuleu, Suchitepequez, Escuintla and Santa Rosa; and Verapaces includes Alta Verapaz and Baja Verapaz.

Table 6
Regression of efficiency ratio (Y/Y^*) on indicators at municipality level.

Coefficient	(1) White Maize	(2)	(3) Black Beans	(4)	(5) Coffee	(6)
Illiteracy rate	−0.169*	−0.168**	−0.079*	−0.050	−0.062	−0.137
	(0.087)	(0.071)	(0.042)	(0.042)	(0.096)	(0.089)
Share of indigenous population	0.232	0.312	0.116	0.199*	0.185**	0.297**
	(0.136)	(0.200)	(0.087)	(0.099)	(0.085)	(0.117)
Road connectivity index	0.170***	0.161**	0.099*	0.103*	0.122	0.047
	(0.057)	(0.073)	(0.052)	(0.053)	(0.079)	(0.089)
Share of households with electricity	−0.039	−0.020	−0.049	−0.039	−0.060	−0.072
	(0.079)	(0.066)	(0.037)	(0.034)	(0.080)	(0.076)
Share of households with cell phones	0.082	0.095	0.006	0.018	−0.002	0.011
	(0.067)	(0.066)	(0.038)	(0.034)	(0.065)	(0.048)
Rate of extortions	−0.532*	−0.474	−0.061	−0.041	−0.238	−0.031
	(0.276)	(0.341)	(0.135)	(0.150)	(0.148)	(0.183)
Constant	−0.126	−0.072	−0.033	0.038	−0.044	−0.150
	(0.185)	(0.191)	(0.087)	(0.108)	(0.164)	(0.149)
Regional fixed effects	No	Yes	No	Yes	No	Yes
Observations	139	139	126	126	108	108
R-squared	0.203	0.227	0.075	0.148	0.100	0.255

Note: Each observation corresponds to a municipality. The efficiency ratios correspond to $\alpha = 0.3$. Variables standardized prior to the regression. Standard errors reported in parentheses clustered at the department level.
*significance at 10% level.
**significance at 5% level.
***significance at 1% level.

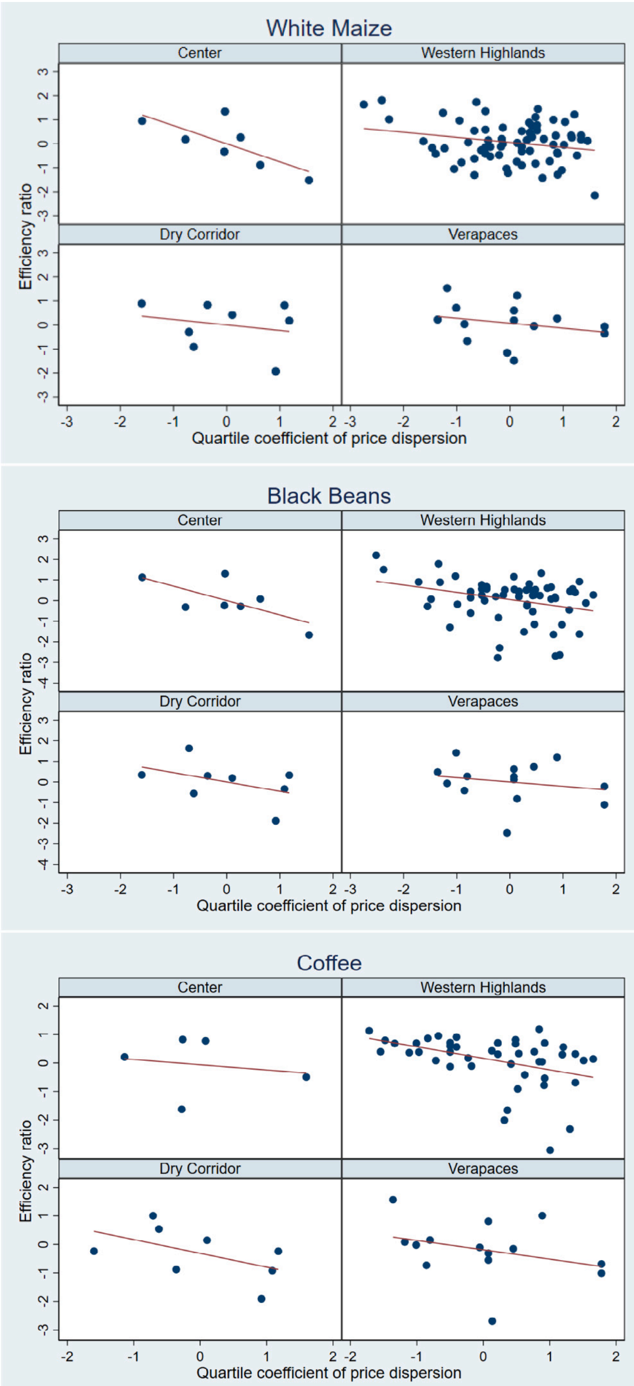

Fig. 3. Efficiency ratio and price dispersion of best land in each municipality. Note: The quartile coefficient of price dispersion is equal to the difference between the 75-th and 25-th percentile of land prices divided by the sum of both percentiles, where price is the price per hectare that surveyed farmers considered to be the most productive agricultural land in their municipality. The efficiency ratios are derived assuming $\alpha = 0.3$. Variables are standardized for comparability purposes across regions within each crop.

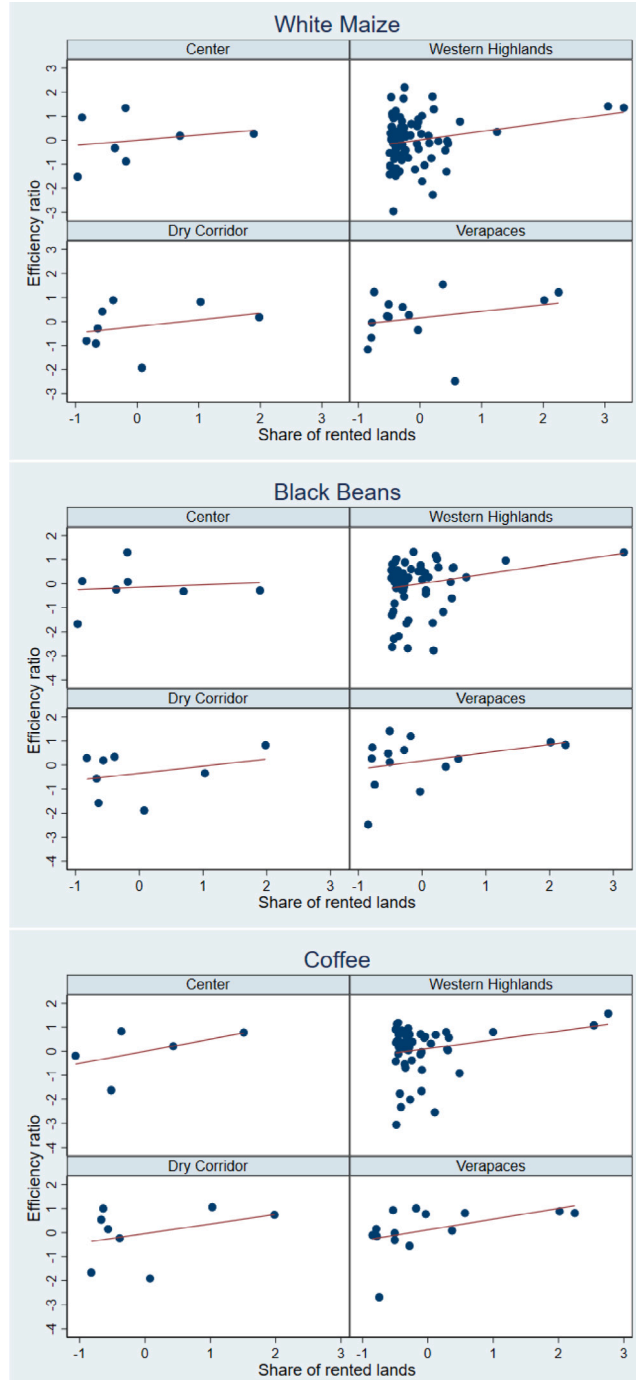

Fig. 4. Efficiency ratio and share of rented lands in each municipality. Note: The efficiency ratios are derived assuming $\alpha = 0.3$. Variables are standardized for comparability purposes across regions within each crop.

maize and black beans than for coffee.[21] Although not reported, we find a positive correlation at the farmer level between the recovered rental prices (based on optimal condition (2)) and our s measure of

[21] We also estimated efficiency ratios for yellow maize, which is less common than white maize, and sugar cane, which is mainly produced in the south of Guatemala. We find that yellow maize shows a similar efficiency ratio than white maize, while sugar cane exhibits an even lower efficiency ratio than coffee. Additional details are available upon request.

farmer productivity for all three crops, which further suggests that the estimated distortions affect relatively more productive farmers.

Note that all efficiency ratios increase when considering a lower α value (0.2) than the reference value of 0.3 and decrease when considering a higher α value (0.4), while the differences across crops and regions remain when working with these alternative values. This is explained by the fact that a larger α value implies a lower level of mechanization or, equivalently, a higher importance of land relative to other factors in the production technology and thereby results, ceteris paribus, in higher distortions from misallocating this factor across farmers. Considering the relatively lower level of mechanization in maize and beans production compared to coffee, the efficiency ratios for $\alpha = 0.4$ could be viewed as an estimated lower bound for these crops (76% on average), whereas the efficiency ratios for $\alpha = 0.2$ could be viewed as an estimated upper bound for coffee (77%).[22]

Table A.4 in the Appendix reports efficiency ratios using land size instead of number of producers in each department as weights. The estimated output gaps and differences across crops and regions are not sensitive to this alternative weighting. The average efficiency ratio is around 82% for maize and beans and 65% for coffee for $\alpha = 0.3$. Peten-Izabal is similarly the region with the largest efficiency ratios for each crop and Pacífico-Bocacosta exhibits one of the lowest ratios.

Lastly, in light of the recent discussion in Gollin and Udry (2021), we recognize that our results could be affected by the presence of other potential unobserved factors such as measurement error not accounted for in the estimations. The authors particularly show that cross-plot measurement error within farms in Tanzania and Uganda appears to be an important source of unobserved variation in productivity that can affect the estimated efficiency gains through land reallocation. While the lack of specific input–output plot-level data prevents us from implementing a similar approach as these authors, we can still assess the sensitivity of our results to excluding multi-plot farmers (52% of our working sample).[23] As shown in Table A.5 in the Appendix, the resulting inefficiencies when only considering the subsample of farmers with one plot are very similar to those when considering the full sample: the average efficiency ratio is 81% for both maize and beans and 66% for coffee for $\alpha = 0.3$. These findings provide additional support to our base results.

4.2. Adjusted benchmark

The benchmark considered above results from a special case in our theoretical framework where all transaction costs beyond the rental cost of land are assumed to vanish. In practice, however, this situation might not be fully achievable as, for example, particular geographic characteristics and cultural factors, which may still generate transaction costs, cannot be eliminated. We accordingly assume that the output gap from our base results in a given department can be reduced up to a certain extent. We consider an adjusted benchmark that consists in first identifying the department in Guatemala with the highest share of rented agricultural land: Escuintla, which is in Pacifico-Bocacosta and is characterized by a larger prevalence of commercial farming. Then, under the assumption that a higher incidence of land rentals in an area is likely associated with lower land market distortions, we derive the ratio of rented land in each department relative to Escuintla. Consequently, if land markets (and thereby the level of distortions) across the country behaved as in Escuintla, the extent to which the output gap could be reduced in each department would be proportional to 1 minus the ratio of rented land relative to Escuintla.[24] This exercise is, in essence, similar to Restuccia and Santaeulalia-Llopis (2017) who compare potential reallocation gains between farms with no marketed land relative to farms with marketed land in Malawi.

Table 5 reports the corresponding output gaps by crop for the baseline and adjusted benchmark.[25] By construction, these adjusted output gaps are lower than the gaps that result from our base findings presented in Section 4.1. We observe that the output gaps, on average, decrease in four percentage points for maize and beans and six percentage points for coffee. This suggests that if land markets across Guatemala were at least as active as in the department with potential less market distortions, there will still be sizeable reallocation gains in the order of 18% for maize and beans and 35% for coffee (compared to 23%, 24% and 46%, respectively, in our base results).

4.3. Output efficiency, price dispersion and rental markets

One of the implications from the theoretical framework described in Section 2 is that we should expect a higher dispersion in land prices among areas with larger market distortions (output inefficiencies). In particular, higher transaction costs τ in an area should result in more dispersed land prices and in a sub-optimal share of land transactions, which prevents the most productive farmers to work at their (larger) optimal scale. To explore this model implication, we rely on a complementary dataset from a three-year panel survey of households collected between 2012 and 2014 over half of the municipalities in the country.[26] The survey included a module on agricultural land markets that inquired about land prices, self-reported quality, and transactions. In one of the questions, households (smallholders) were asked to provide the price per hectare of what they would consider to be the most productive agricultural land in their municipality, which permits us to derive a measure of price dispersion at the municipality level.

Fig. 3 plots the corresponding efficiency ratios (Y/Y^*) based on the distribution of farmers' productivities estimated at the municipality level from the census microdata (for $\alpha = 0.3$) against the quartile coefficient of dispersion (QCD) of the reported land prices for the municipalities covered in the supplementary survey.[27] We observe an inverse efficiency-price dispersion relationship among all three crops and regions, in line with the theoretical implication that more efficient areas (municipalities) should show a lower dispersion in land prices.[28] Fig. 4 presents, in turn, scatterplots of efficiency ratios and the share of agricultural land that is reported rented in a municipality. We observe a positive relationship between these two variables suggesting that those

[22] In Table A.3 in the Appendix we also report the results when working at the municipality level; that is, when farmers possible area of operation is within their municipality (instead of their department). The aggregate efficiency ratios derived based on the distribution of farmers' productivities at the municipality level are, on average, marginally higher than at the department level (reported in Table 4): 82% for maize, 81% for beans, and 71% for coffee for $\alpha = 0.3$. In general, the smaller the geographic area of operation (i.e., the possible area to rent in/out land), the smaller the potential distortions as transaction costs become arguably lower due to the smaller market size, and vice versa.

[23] We assume that the amount of land in a given department is determined by the sum of land of all farmers reporting one plot, which can rent in/out land within their department.

[24] The adjusted output gap is equal to $(1-(Y/Y^*))(1-R_e)/[(Y/Y^*)+(1-(Y/Y^*))(1-R_e)]$, where Y/Y^* is the estimated output efficiency ratio from our base results and R_e is the ratio of share rented land in a given department compared to the share in Escuintla.

[25] For ease of exposition, hereafter we present results for $\alpha = 0.3$.

[26] The survey was part of the evaluation of a large-scale program executed by the Government of Guatemala against food insecurity and malnutrition and covered 176 of the 340 municipalities in the country, particularly the poorest and with the highest stunting rates.

[27] The QCD is equal to the difference between the 75-th and 25-th percentile of prices divided by the sum of both percentiles. In the plots we only include municipalities with at least ten price observations in the supplementary survey.

[28] The same pattern holds if we use instead the coefficient of price variation or if we use the price per hectare that the farmer valued her own land, after controlling for self-reported land quality (on a scale 1-10).

municipalities with a higher prevalence of rented land are also those with seemingly lower land market distortions.

Overall, these plotted relationships at the municipality level provide suggestive evidence that more efficient areas exhibit both lower dispersion in land prices and more active rental markets. The reduced number of data points prevents us, however, from deriving formal pairwise correlations (slopes) with statistical precision, except for the Western Highlands where the estimated correlations are in the range of −0.24 to −0.41 in Fig. 3 and 0.22 to 0.23 in Fig. 4, which are statistically significant at conventional levels.[29]

4.4. Potential channels of distortions

We now turn to examine whether the aggregate output inefficiencies co-variate with certain observable characteristics to identify potential channels that could be driving the estimated distortions. For this purpose, we regress the efficiency ratios (Y/Y^*) derived at the municipality level by crop on a set of indicators related to education, ethnicity, road and services accessibility, and level of insecurity in the area, obtained from multiple data sources and different available years. These indicators include illiteracy rate, rate of indigenous population, road connectivity index, share of households with electricity and cellphones, and rate of extortions.[30] Table 6 presents the estimation results. The variables were first standardized for comparability purposes and the reported standard errors are clustered at the department level.[31] Columns (2), (4) and (6) include regional fixed effects.

We find that road accessibility is positively associated with the estimated efficiency ratios, particularly for maize and beans; the more connected the municipality in terms of paved (and unpaved) roads, normalized by extension area and population, the higher the efficiency ratio. Hence, higher transaction costs resulting from lower accessibility could be playing some role in explaining land market (output) inefficiencies in an area. Similarly, we observe a positive correlation between ethnicity and efficiency, especially in coffee and beans; the larger the share of indigenous population in a municipality, the higher the efficiency ratio, which suggests that cultural aspects may also be explaining part of the observed distortions. This correlation could be linked with the fact that we typically observe more social cohesion within rural areas dominated by indigenous populations in Latin America (CEPAL, 2007): there is more trust and less information asymmetries; in contrast, in rural areas with a lower prevalence of indigenous populations, there are probably more cultural barriers and information asymmetries between neighboring populations. Illiteracy rate is, in turn, negatively correlated with output efficiency for maize and beans; all else equal, we expect a higher market dynamism and subsequent better land allocation among locations with more educated people.

Regarding the other indicators, the rate of extortions, which captures the level of insecurity (conflict) in an area, is negatively associated with output efficiency, but the estimated partial correlations are only marginally significant for maize. Access to services such as electricity and cellphones, which act as proxies of easiness of information flow and development in an area, are not correlated with output efficiency.

[29] The correlations are significant at the 95% level except for beans and coffee in Fig. 4 that are significant at the 90% level.

[30] The illiteracy rate is obtained from the Comision Nacional de Alfabetizacion (CONALFA) for 2014; rate of indigenous population from the Population Census (INE) for 2002; road connectivity index (weighted sum of paved and unpaved road kilometers normalized by extension area and population) from the Ministry of Agriculture (MAGA) for 2008; share of households with electricity and cellphones from the National Survey of Living Conditions (ENCOVI) for 2006; and rate of extortions from INFOSEGURA-Guatemala database for 2017.

[31] The number of observations differs across crops as not all crops are produced in all municipalities and the set of available indicators differs across locations.

5. Concluding remarks

Farm size and land allocation play an important role in explaining lagging agricultural productivity in developing countries. This paper assesses the impact of land market distortions on land allocation and aggregate agricultural productivity. We develop a theoretical model to examine to what extent market distortions can explain non-optimal land allocation and output efficiency. We then quantify these distortions using census microdata from Guatemala. The estimation results show that due to land market imperfections aggregate output is roughly 81% of the efficient output for both white maize and black beans and 69% for coffee, which are the three major crops produced across the country. More efficient areas seem to exhibit a lower dispersion in land prices and more dynamic rental markets, in line with our theoretical discussion. We also observe that the degree of land market distortions across areas co-variate to some extent with road accessibility, ethnicity, and with the level of education in the area.

While there are some variations in efficiencies across regions, the overall findings indicate the presence of larger distortions for high-value export crops such as coffee (with an estimated output gap of around one third), relative to staple crops such as maize and beans (with an output gap of around one fifth). These results suggest that the latter crops, with a higher share of small-scale and subsistence agriculture, may already be operating close to their maximum production potential such that eliminating land market distortions will have a smaller effect on reaching their optimal output level. In contrast, the elimination of land distortions could have a larger effect on the aggregate productivity of coffee and possibly on other similar high-value cash crops.

The analysis examining potential factors associated with output inefficiencies suggest, for example, the importance of continuing improving accessibility and education as well as further recognizing and addressing likely cultural barriers. Certainly, policies in this regard, such as investment in road infrastructure and education, will require some time to become effective, while overcoming cultural differences may be more difficult. From the analysis, areas with a higher prevalence of indigenous population already seem to be operating more efficiently. Considering that the mobile penetration rate in rural Guatemala is over 90%, market information systems exploiting new technologies of information could also help, at least in the short-term, to develop or expand rental land markets across the country, maybe within areas that share similar cultural (ethnic) characteristics, and reallocate land from less to more productive producers. Pilot programs to assess farmers' willingness to rent in/out land and whether providing market information effectively contributes to the generation of rental markets are avenues of future work along these lines.

Finally, while several robustness checks support our main findings, we recognize that we cannot fully discard the presence of potential measurement error and other unobservables (such as input quality) in our estimations. Data restrictions also require to follow a two-stage approach to derive our productivity measure. Our results should thus be interpreted with caution given the nature of our data. Similarly, the analysis is based on data from the 2003 agricultural census, while a more recent census in the country is still lacking. If, for instance, production technologies have generally improved in Guatemala over the past years, our estimation approach would imply lower gains from an efficient allocation of land, holding constant other factors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

See Figs. A.1–A.2 and Tables A.1–A.5.

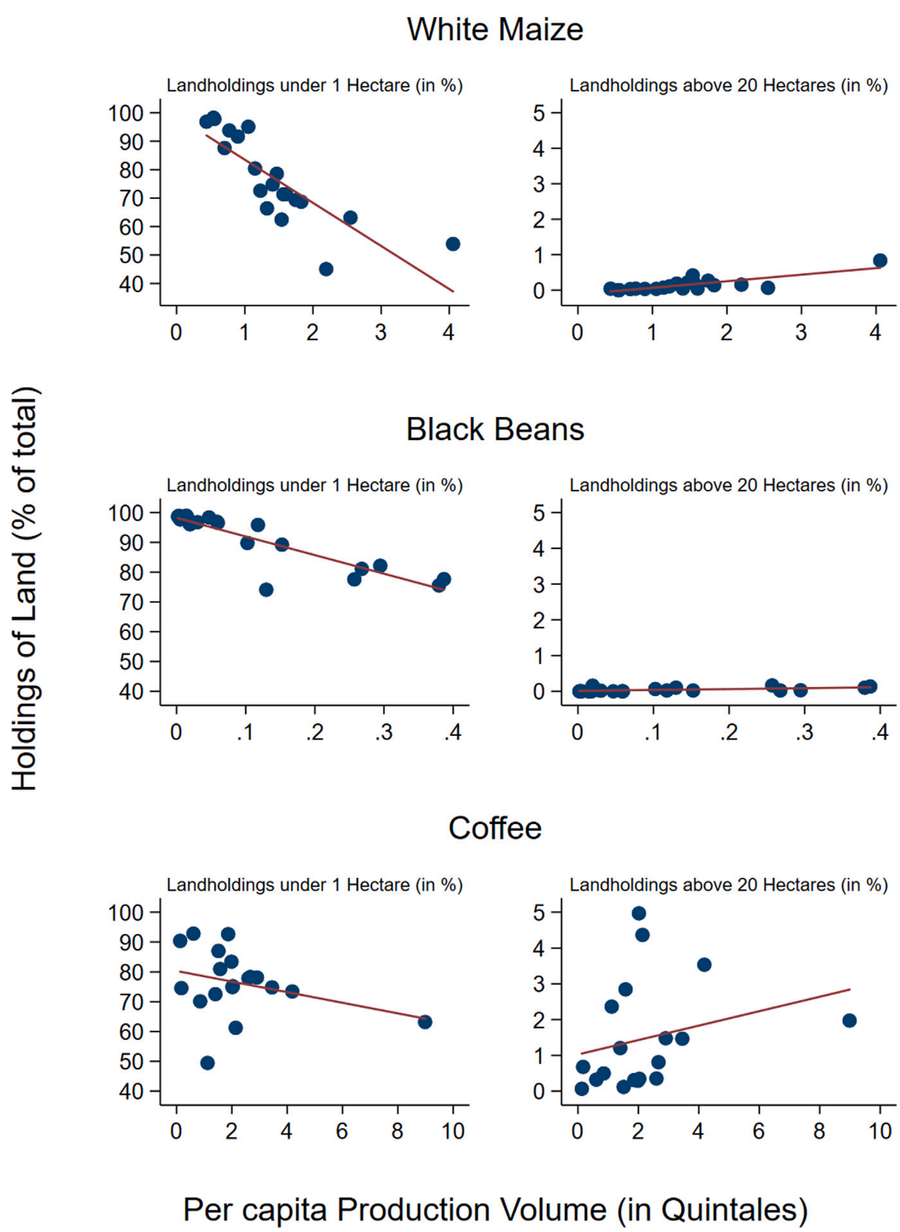

Fig. A.1. Share of small and very large landholdings and per capita production volume across departments by crop. Note: Landholding size is defined as the number of hectares dedicated to the production of each crop. The volume of production is measured in Guatemalan quintales where one quintal is equivalent to 100 pounds.

Table A.1
Self-reported land quality dispersion and land price variance at the municipality versus department and regional level.

Panel A: Land quality dispersion at the municipality level				
	Mean	Min	Max	Number of Municipalities
Standard Deviation	1.64	0	3.40	141
Coefficient of Variation	0.26	0	0.59	141

Panel B: Percentage of municipalities where land price variance at department or regional level is greater or equal than at municipality level			
	Land Prices		Number of Municipalities
	Best owned land	Best land in Municipality	
If Departmental ≥ Municipal	60.9%	63.6%	110
If Regional ≥ Municipal	69.1%	71.8%	110

Note: In Panel A, land quality is measured from 1 to 10 where farmers in each municipality are asked "*How would you rate your best land on a scale from 1 to 10 (where 10 is equivalent to the most productive land in your municipality)?*". The values reported correspond to the mean, minimum, and maximum standard deviation and coefficient of variation across municipalities. In Panel B, the values reported represent the percentage of municipalities where the departmental (for the first row, regional for the second) land price variance is larger than the municipal land price variance. The first and second column indicate the price variable used to calculate the variance: farmers' reported price for their highest-quality land in the first column, and farmers' reported price for the highest-quality land available in their municipality in the second column.

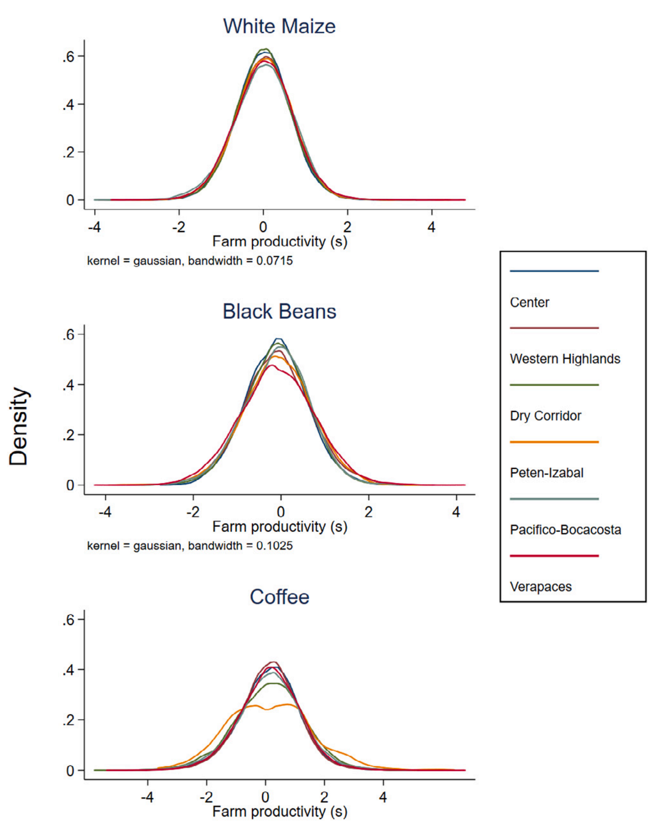

Fig. A.2. Distribution of estimated farmers productivity by region. Note: Farmers productivity derived from the full-sample estimations depicted in Table A.2.

Table A.2
Regression of derived Ln s measure on set of characteristics at farmer level, full sample.

Coefficient	(1) White Maize	(2) Black Beans	(3) Coffee
Age	0.001*	−0.002***	0.004***
	(0.000)	(0.000)	(0.001)
If female	−0.198***	−0.173***	−0.263***
	(0.009)	(0.017)	(0.023)
Years of schooling	0.003*	0.007**	0.039***
	(0.001)	(0.003)	(0.006)
Household Labor/Total Labor	0.629***	0.642***	0.148
	(0.072)	(0.045)	(0.133)
If has machinery	0.126***	0.008	0.178***
	(0.028)	(0.018)	(0.037)
If has equipment	0.074***	0.050***	0.127***
	(0.015)	(0.009)	(0.028)
If has livestock	0.041**	0.006	−0.058**
	(0.016)	(0.034)	(0.021)
If uses high-performance seeds	0.069***	0.074***	−0.003
	(0.011)	(0.024)	(0.045)
If uses organic fertilizer	0.003	0.064***	0.130***
	(0.010)	(0.022)	(0.027)
If uses chemical fertilizer	0.062**	0.045	0.080*
	(0.029)	(0.047)	(0.042)
If uses pesticide	0.101***	0.019	−0.018
	(0.015)	(0.028)	(0.026)
If has irrigation system	0.071**	0.089*	0.141*
	(0.033)	(0.048)	(0.075)
Number of different crops	−0.309***	−0.250***	−0.391***
	(0.014)	(0.015)	(0.016)
Constant	2.283***	1.510***	1.516***
	(0.035)	(0.082)	(0.088)
Municipality fixed effects	Yes	Yes	Yes
Observations	396,317	113,133	147,353
R^2	0.456	0.485	0.380

Note: The Ln s measure used as the dependent variable in the regressions is derived assuming $\alpha = 0.3$. The regressions include an indicator variable that takes the value of one for farmers with missing gender, age, and years of education, and zero otherwise. Standard errors reported in parentheses clustered at the department level.
*Significance at 10% level.
**Significance at 5% level.
***Significance at 1% level.

Table A.3
Efficiency ratios (Y/Y^*), based on calculations at Municipality level.

Region	White Maize			Black Beans			Coffee		
	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$
All regions	87.6%	82.2%	77.1%	86.8%	81.3%	76.1%	78.9%	70.8%	63.7%
Center	87.9%	82.7%	77.7%	87.5%	82.2%	77.3%	74.6%	66.5%	60.4%
Western Highlands	87.6%	82.3%	77.3%	85.9%	80.2%	74.9%	79.6%	71.7%	64.6%
Dry Corridor	87.3%	81.6%	76.3%	87.3%	81.7%	76.5%	78.4%	69.8%	63.1%
Peten-Izabal	89.2%	84.4%	79.8%	88.8%	83.9%	79.4%	79.0%	70.8%	63.7%
Pacifico-Bocacosta	86.8%	81.3%	76.2%	85.8%	80.1%	75.0%	74.3%	65.7%	58.9%
Verapaces	87.4%	81.8%	76.3%	86.5%	80.7%	75.3%	80.1%	71.9%	64.4%

Note: The regional ratios reported are the corresponding averages of the municipality ratios in a region weighted by the number of producers in each municipality. The ratio for All regions is the weighted average across regions. The Center region includes the departments of Sacatepequez and Chimaltenango; Western Highlands includes Huehuetenango, Quiche, San Marcos, Quetzaltenango, Totonicapan and Solola; Dry Corridor includes Chiquimula, Jutiapa, Jalapa, El Progreso and Zacapa; Peten-Izabal includes Peten and Izabal; Pacifico-Bocacosta includes Retalhuleu, Suchitepequez, Escuintla and Santa Rosa; and Verapaces includes Alta Verapaz and Baja Verapaz.

Table A.4
Efficiency ratios (Y/Y^*), weighted by land size.

Region	White Maize			Black Beans			Coffee		
	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$
All regions	87.1%	81.6%	76.4%	87.2%	81.7%	76.6%	74.6%	65.4%	57.8%
Center	87.5%	82.0%	76.9%	87.8%	82.4%	77.4%	69.7%	59.7%	52.2%
Western Highlands	86.1%	80.3%	75.0%	86.0%	80.1%	74.6%	73.8%	64.5%	56.8%
Dry Corridor	86.8%	81.0%	75.5%	86.1%	80.2%	74.7%	78.0%	69.4%	61.8%
Peten-Izabal	88.7%	83.7%	79.0%	88.5%	83.5%	78.8%	79.9%	71.6%	63.9%
Pacifico-Bocacosta	84.5%	78.4%	73.2%	86.2%	80.3%	74.9%	73.7%	64.4%	57.0%
Verapaces	87.2%	81.5%	76.0%	85.8%	79.7%	74.1%	78.6%	70.0%	62.5%

Note: The regional ratios reported are the corresponding averages of the departmental ratios in a region weighted by land size in each department. The ratio for All regions is the weighted average across regions. The Center region includes the departments of Sacatepequez and Chimaltenango; Western Highlands includes Huehuetenango, Quiche, San Marcos, Quetzaltenango, Totonicapan and Solola; Dry Corridor includes Chiquimula, Jutiapa, Jalapa, El Progreso and Zacapa; Peten-Izabal includes Peten and Izabal; Pacifico-Bocacosta includes Retalhuleu, Suchitepequez, Escuintla and Santa Rosa; and Verapaces includes Alta Verapaz and Baja Verapaz.

Table A.5
Efficiency ratios (Y/Y^*), considering only farmers that report one plot.

Region	White Maize			Black Beans			Coffee		
	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$	$\alpha = .2$	$\alpha = .3$	$\alpha = .4$
All regions	86.9%	81.3%	76.1%	86.6%	80.8%	75.4%	75.3%	66.4%	59.0%
Center	88.5%	83.3%	78.4%	88.3%	83.1%	78.1%	69.0%	59.7%	52.9%
Western Highlands	86.4%	80.7%	75.6%	86.0%	80.0%	74.5%	77.3%	68.8%	61.7%
Dry Corridor	87.2%	81.6%	76.1%	86.0%	80.1%	74.5%	75.8%	66.7%	58.9%
Peten-Izabal	88.7%	83.6%	78.7%	88.8%	83.8%	79.0%	76.9%	67.2%	58.3%
Pacifico-Bocacosta	84.3%	78.3%	73.1%	84.0%	77.6%	72.0%	70.2%	60.2%	52.5%
Verapaces	87.3%	81.6%	76.1%	86.0%	80.0%	74.4%	75.5%	66.3%	58.7%

Note: The regional ratios reported are the corresponding averages of the departmental ratios in a region weighted by the number of producers in each department. The ratio for All regions is the weighted average across regions. The Center region includes the departments of Sacatepequez and Chimaltenango; Western Highlands includes Huehuetenango, Quiche, San Marcos, Quetzaltenango, Totonicapan and Solola; Dry Corridor includes Chiquimula, Jutiapa, Jalapa, El Progreso and Zacapa; Peten-Izabal includes Peten and Izabal; Pacifico-Bocacosta includes Retalhuleu, Suchitepequez, Escuintla and Santa Rosa; and Verapaces includes Alta Verapaz and Baja Verapaz.

References

Adamopolous, T., Restuccia, D., 2014. The size distribution of farms and international productivity differences. Amer. Econ. Rev. 104 (6), 1667–1697.

Adamopoulos, T., Restuccia, D., 2017. Geography and Agricultural Productivity: Cross-country Evidence from Micro Plot-Level Data. 2017 Meeting Papers No. 1180, Society for Economic Dynamics.

Aragon, F., Restuccia, D., Rud, J.P., 2019. Are small farms really more productive than large farms? NBER Working Paper No. 26331.

Bardales, W.A., Castañon, C., Herrera, J.L., 2019. Clima de Guatemala, tendencias observadas e indices de cambio climatico. In: Castellanos, E.J., Paiz-Estevez, A., Escriba, J., Rosales-Alconero, M., Santizo, A. (Eds.), Primer reporte de evaluacion del conocimiento sobre cambio climatico en Guatemala. Editorial Universitaria UVG, Guatemala, pp. 20–39 (in Spanish).

Barrett, C.B., 1996. On price risk and the inverse farm size-productivity relationship. J. Dev. Econ. 51 (2), 193–215.

Barrett, C.B., Bellemare, M.F., Hou, J.Y., 2010. Reconsidering conventional explanations of the inverse productivity–size relationship. World Dev. 38 (1), 88–97.

Bartelsman, E., Haltiwanger, J., Scarpetta, S., 2013. Cross-country differences in productivity: The role of allocation and selection. Amer. Econ. Rev. 103 (1), 305–334.

Bento, P., Restuccia, D., 2017. Misallocation, establishment size, and productivity. Am. Econ. J.: Macroeconomics 9 (3), 267–303.

Besley, T.J., Leight, J., Pande, R., Rao, V., 2012. The regulation of land markets: Evidence from tenancy reform in India. CEPR Discussion Paper No. 8736.

BID, 2019. Perfil de riesgo de desastres por sequía: El Salvador, Guatemala y Honduras. Nota Tecnica IDB-TN-01888, July, Banco Interamericano de Desarrollo, Washington DC (in Spanish).

Carrera, J.A., 1999. El estudio de mercado de tierras en Guatemala. In: Serie de Desarrollo Productivo N. 73. CEPAL/GTZ: Santiago de Chile.

Caselli, F., 2005. Accounting for cross-country income differences. In: Handbook of Economic Growth, Vol. 1. Elsevier, pp. 679–741.

CEPAL, 2007. Cohesion social: Inclusion y sentido de pertenencia en America Latina y el Caribe. Naciones Unidas, Chile, Santiago de Chile (in Spanish).

Chamberlain, J., Ricker-Gilbert, J., 2016. Participation in rural land rental markets in Sub-Saharan Africa: Who benefits and by how much? Evidence from Malawi and Zambia. Am. J. Agric. Econ. 98 (5), 1507–1528.

Chen, C., 2017. Untitled land, occupational choice, and agricultural productivity. Am. Econ. J.: Macroeconomics 9 (4), 91–121.

Chen, C., Restuccia, D., Santaeulália-Llopis, R., 2021. The effects of land markets on resource allocation and agricultural productivity. Rev. Econ. Dyn. (ISSN: 1094-2025) http://dx.doi.org/10.1016/j.red.2021.04.006.

David, J.M., Hopenhayn, H.A., Venkateswaran, V., 2016. Information, misallocation, and aggregate productivity. Q. J. Econ. 131 (2), 943–1005.

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

Durr, J., 2016. The political economy of agriculture for development today: The "small versus large" scale debate revisited. Agric. Econ. 47, 671–681.

Foster, A.D., Rosenzweig, M.R., 2017. Are there too many farms in the world? Labor-market transaction costs, machine capacities and optimal farm size. NBER Working Paper No. 23909.

Gauster, S., Isakson, R., 2007. Eliminating market distortions, perpetuating rural inequality: An evaluation of market-assisted land reform in Guatemala. Third World Q. 28 (8), 1519–1526.

Goldstein, M., Udry, C., 2008. The profits of power: Land rights and agricultural investment in Ghana. J. Polit. Econ. 116 (6), 981–1022.

Gollin, D., Lagakos, D., Waugh, M.E., 2014. The agricultural productivity gap. Q. J. Econ. 129 (2), 939–993.

Gollin, D., Udry, C., 2021. Heterogeneity, measurement error and misallocation: Evidence from African agriculture. J. Polit. Econ. 129 (1), 1–80.

Gould, K.A., 2006. Land regularization on agricultural frontiers: The case of Northwestern Petén, Guatemala. Land Use Policy 23 (4), 395–407.

Henderson, H., Isaac, A.G., 2017. Modern value chains and the organization of agrarian production. Am. J. Agric. Econ. 99 (2), 379–400.

Hopenhayn, H., 1992. Entry, exit, and firm dynamics in long run equilibrium. Econometrica 60 (5), 1127–1150.

Hopenhayn, H., Rogerson, R., 1993. Job turnover and policy evaluation: A general equilibrium analysis. J. Polit. Econ. 101 (5), 915–938.

Hsieh, C.-T., Klenow, P.J., 2009. Misallocation and manufacturing TFP in China and India. Q. J. Econ. 124 (4), 1403–1448.

INE, 2006. Encuesta Nacional Agropecuaria (ENA) 2006. Instituto Nacional de Estadistica, Guatemala, Ciudad de Guatemala (in Spanish).

Jayne, T.S., Chamberlin, J., Traub, L., Sitko, N., Muyanga, M., Yeboah, F.K., Anseeuw, W., Chapoto, A., Wineman, A., Nkonde, C., et al., 2016. Africa's changing farm size distribution patterns: The rise of medium-scale farms. Agric. Econ. 47 (S1), 197–214.

Lagakos, D., Waugh, M.E., 2013. Selection, agriculture and cross-country productivity differences. Amer. Econ. Rev. 103 (2), 948–980.

Lastarria-Cornhiel, S., 2003. Guatemala Country Brief: Property Rights and Land Markets. Land Tenure Center, University of Wisconsin-Madison.

Lowder, S.K., Sanchez, M.V., Bertini, R., 2019. Farms, Family Farms, Farmland Distribution and Farm Labour: What Do We Know Today? FAO Agricultural Development Economics Working Paper 19-08, FAO, Rome.

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

MAGA, 2011. El Agro en Cifras 2011. Ministerio de Agricultura, Ganaderia y Alimentacion, Guatemala, Ciudad de Guatemala (in Spanish).

MAGA, 2013. El Agro en Cifras 2013. Ministerio de Agricultura, Ganaderia y Alimentacion, Guatemala, Ciudad de Guatemala (in Spanish).

Place, F., 2009. Land tenure and agricultural productivity in Africa: A comparative analysis of the economics literature and recent policy strategies and reforms. World Dev. 37 (8), 1326–1336.

Restuccia, D., Rogerson, R., 2008. Policy distortions and aggregate productivity with heterogeneous plants. Rev. Econ. Dyn. 11 (4), 707–720.

Restuccia, D., Rogerson, R., 2017. The causes and costs of misallocation. J. Econ. Perspect. 31 (3), 151–174.

Restuccia, D., Santaeulalia-Llopis, R., 2017. Land misallocation and productivity. NBER Working Paper No. 23128.

Restuccia, D., Yang, D.T., Zhu, X., 2008. Agriculture and aggregate productivity: A quantitative cross-country analysis. J. Monetary Econ. 55 (2), 234–250.

Valentinyi, A., Herrendorf, B., 2008. Measuring factor income shares at the sectoral level. Rev. Econ. Dyn. 11 (4), 820–835.