THREE CASE STUDIES ON THE IMPACT OF INEQUALITY ON AGRICULTURAL EFFICIENCY, PRODUCTIVITY, AND THE ADOPTION AND INTENSITY OF USE OF CLIMATE-SMART PRACTICES

A Dissertation presented to
the Division of Applied Social Sciences
at the University of Missouri-Columbia

In partial fulfilment
of the requirements for the degree
Doctor of Philosophy

---

by
HERNÁN BORRERO
Supervisor
DR. THEODOROS SKEVAS
MAY 2024

APPROVAL PAGE

The undersigned, appointed by the dean of the Graduate School, have examined the Dissertation entitled

THREE CASE STUDIES ON THE IMPACT OF INEQUALITY ON AGRICULTURAL EFFICIENCY, PRODUCTIVITY, AND THE ADOPTION AND INTENSITY OF USE OF CLIMATE-SMART PRACTICES

presented by Hernan Borrero, a candidate for the degree of Doctor of Philosophy in Agricultural & Applied Economics, and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Theodoros Skevas

Dr. Harvey James

Dr. Laura McCann

Dr. Shawn Ni

ACKNOWLEDGEMENTS

First of all, I wish to thank Dr. Theodoros Skevas for believing in me since the beginning; for helping me to navigate the difficult waters of Ph.D. studies (ultimately, bringing me to safe port); and for making me a much more confident economist. Second, I wish to thank each of the members of my Doctoral Committee: Dr. Harvey James, for accepting me into the program in the first place and for supporting me throughout, especially in the face of difficulty; Dr. Laura McCann, for teaching me how to hold amenable and fruitful academic discussions, as I hope to have demonstrated in the defense; and Dr. Ni, for teaching me how to be a much better writer. I recommend many classes in this program; but his 'research workshop' is a must for everyone doing graduate studies in economics and related fields. I also want to thank Dr. Corinne Valdivia, for accompanying me at the beginning of this process; and Dr. Wang, for the multiple coffee conversations. All of them have made me a much better professional and, most importantly, a much better person.

TABLE OF CONTENTS

ABSTRACT

In this three-chapter Dissertation we explore the effect of inequality—particularly *land* inequality—on three different aspects of agricultural production: technical inefficiency, total factor productivity, and the adoption and intensity of use of climate-smart agricultural practices. In the first chapter, we employ a sample of fifteen Latin American rural sectors over a period of seventeen years (2005-2021), and implement a Stochastic Frontier Analysis methodology, finding a positive effect of inequality on technical inefficiency. In the second chapter, we use a sample of 299 transitory-crop farmers of the Peruvian highlands and two alternative empirical strategies: a dynamic panel data model, and a control function approach; finding a negative effect of land inequality (measured by the Gini and Theil coefficients) on total factor productivity. And finally, in the third chapter, we exploit a larger sample of over four-thousand farmers from the same Peruvian region and a double-hurdle modelling approach, finding a negative effect of land inequality (measured by the Gini index and the Kakwani relative land deprivation index) on the probability of adoption of climate-smart agricultural practices.

GENERAL INTRODUCTION

In this three-essay Dissertation, we explore the effect of economic inequality—and particularly, *land* inequality—on three different aspects of agricultural production: technical inefficiency, total factor productivity, and adoption of climate-smart agricultural practices.

The point of departure for establishing a relationship between inequality and these three aspects of agricultural production is the agricultural economics theoretical framework of factor market imperfections. The fundamental economic idea of this framework is that land, capital, and labor market imperfections in developing-country agricultural contexts cause small and large farmers to face different shadow prices for these factors of production. In particular, while small farms face a relatively low opportunity cost for labor (due to their heavy reliance on family labor) and high shadow prices for capital and land (due to credit rationing, covariate risk, land speculation, and other well-documented market failures), large farms face relatively high costs for labor (due to monitoring costs) and relatively low shadow prices for land and capital.

Of course, if it weren't for factor market imperfections, all factor prices would converge to a unique price reflecting their true social opportunity costs. However, the existence of these imperfections is widely documented in the agricultural economics literature. The academic task of this dissertation has been, then, to elicit the theoretical consequences of this divergence for agricultural production, and to test whether the predicted outcomes are indeed observed in our contexts of study. Specifically, we propose three possible repercussions:

- First, that the lower shadow prices for land and capital faced by large farmers leads them to treat land as a relatively abundant resource, even in land-scarce economies, and to substitute machinery for labor, even in labor-abundant economies. As a result, there is a socially inefficient allocation of resources, thus leading to macroeconomic *technical inefficiency*.

- Second, that large agricultural farms, on average, use land more extensively, thus reducing their per-hectare productivity. Nevertheless, it is not sufficient to test whether land inequality affects yield, since land productivity reductions could potentially be compensated by productivity increases in labor and machinery. Hence, a total factor productivity approach is needed to fully understand the inequality-productivity relationship.
- And third, because small farms face a relatively low shadow price for labor and many climate-smart agricultural practices tend to be very labor-intensive (e.g., application of organic matter, complex intercropping systems, etc.), it is therefore more likely that small farms will adopt them and use them more intensively. Therefore, a re-distribution of land from the large to the small farms is justified on the grounds of climate adaptation and mitigation of green-house gas emissions, leading to the concept of a climate-smart rural reform.

Putting these three ideas together gives rise to this Dissertation. The first one (the inequality-technical inefficiency relationship) is tested in the context of fifteen Latin American rural sectors during the 2005–2021 period. The choice of this case study is not fortuitous: land in this region remains 'locked-up' in a latifundia-minifundia social structure which can be traced back to the colonial institutions of the *Mita*, the *Encomienda*, and the *composiciones de tierras*, which merged the rights over the land with the paternalistic control of the 'embedded' rural populations. As a result, land played—and continues to play—a fundamental role as an element of power and prestige, rather than as a simple factor of production. As a consequence, we argue, land and rural income inequalities are determined by very-long run processes of institutional formation and other non-productive motivations, quite independent from the short-run vicissitudes of the production process itself, thus allowing us to draw statistical inferences in the empirical exercise that follows, based on Stochastic Frontier Analysis. Overall, we find a significant *positive* effect of rural income inequality on technical *in*efficiency, supporting the agricultural economics perspective.

In the second essay, which looks at the land inequality-productivity relationship, we investigate the effect of *land* inequality on agricultural total factor productivity (TFP) using a sample of transitory crop farms from the Peruvian highlands, over a period of four years (2016–2019). We employ three different measures of land inequality—the Gini, the Theil, and the Kakwani indexes—and two alternative estimation methods: a dynamic panel data model and a control function approach, both addressing input endogeneity and selection bias. Our results show that both the Gini and Theil indexes exert a significant negative impact on TFP—indicating the detrimental impact of land inequality on productivity—but no significant effect of the Kakwani index. Furthermore, we find that there was no significant change in the TFP of the sampled farms over the period, emphasizing the persistent nature of the observed productivity levels.

And finally, we investigate the effect of land inequality on the probability of adoption and intensity of use of climate-smart agricultural practices, yet again, in the Peruvian highland context. For that purpose, we assemble a sample of over four-thousand farms of the region drawn from the 2022 national agricultural survey of Peru, and implement a double-hurdle regression model: in the first hurdle, the farmer decides whether to adopt or not any positive quantity of climate-smart agricultural practices (a total of nine practices were identified, such as crop rotation, intercropping, application of organic matter, terracing, and so on); if this hurdle is ‘crossed’, then he or she decides on the total number of practices to implement—the intensity of use decision. In general, we find a positive effect of land inequality (as measured by the Gini and Kakwani indexes) on the adoption decision, lending support to the idea of a climate-smart rural reform.

In all three cases we have procured to follow the scientific method in that normative values of inequality have been avoided, and in that only the positive elements of the agricultural economics theory and practice guided our efforts. Also, comments that could not be directly substantiated from our findings were avoided to rule out any form of speculation on our part. We hope that this Dissertation is of the interest of the general public and that it contributes meaningfully to the development of the science.

Chapter 1

THE EFFECT OF INEQUALITY ON TECHNICAL INEFFICIENCY: EVIDENCE FROM LATIN AMERICAN RURAL SECTORS

1.1. Introduction

In this paper we examine the neglected relationship between inequality and *technical inefficiency*. This link not only represents an additional pathway through which inequality may affect economic growth and development, but also holds significant relevance for developing-country agricultural sectors; where, as we shall see, the theory predicts a negative effect of inequality on social and economic efficiency.

To investigate this relationship, we use a sample of fifteen Latin American rural sectors over a period of seventeen years (2005–2021). In this region, land remains locked-up in a rigid *latifundia-minifundia* social structure, whose origins can be traced back to the colonial institutions of the *Mita*, the *Encomienda* and the *composiciones de tierras*, among others (García, 1967; Chonchol, 1996). At turn, these historical patterns of land distribution determine current levels of rural income inequality (Sokoloff & Engerman, 2000; Frankema, 2005, 2008), allowing us to draw causal statistical inferences in the empirical part of this work.

Now then, our empirical exercise utilizes stochastic frontier analysis (SFA), a parametric technique that uses output and input information from a sample of productive units (e.g., firms, agricultural sectors, etc.) to measure the gap between the observed and the potential output that could be achieved from the available resources and technology—i.e., the level of technical inefficiency—while allowing to test for its determining factors (Battese & Coelli, 1995; Coelli et al., 2005). Whereas most of the SFA literature focuses on determinants such as human capital and management practices (Sickles & Zelenyuk, 2019), we exploit this methodological framework to investigate the impact of inequality on technical inefficiency. Overall, we find strong support for the outlined theory.

This paper is organized as follows: the next section, Section 1.2, establishes the theoretical link between inequality and technical inefficiency based on a review of the agricultural economics literature; Section 1.3 introduces the SFA methodology; Section 1.4 presents the Latin American case study, divided into three subsections: first, an exposition of the latifundia-minifundia social structure; second, a description of the data and sources; and third, our empirical specification; Section 1.5 presents the results; and lastly, Section 1.6 concludes.

1.2. Literature Review

Since Kuznets (1955), the majority of the literature linking inequality with economic performance focuses on long-run economic *growth* (for a complete overview of the inequality-growth literature refer to Neves & Silva, 2014). More recently, an interesting stream of research has emerged analysing the effects of inequality on very long-run economic development from a *historical* perspective (Sokoloff & Engerman, 2000; Banerjee & Iyer, 2005; Adamopoulos, 2007). However, the link between inequality and *technical inefficiency* remains practically unexplored[1]. Yet, this

---

[1] To the best of our knowledge, only one study looks at this relationship in the specific context of South Africa (Espoir & Ngepah, 2020).

could be an important channel through which inequality may affect long- and very long-run economic outcomes, especially in the context of developing-country agriculture.

Effectively, in the agricultural economics literature, the theory predicts a negative effect of inequality on aggregate (i.e., macroeconomic) technical efficiency. The key element of the explanation are factor-market imperfections, which cause small and large farms in unequal or *bimodal* agrarian systems to face different factor prices (Berry & Cline, 1979; Carter, 1984; Ellis, 1993). Specifically, on one side, small farmers confront a relatively low opportunity cost of labor due to their reliance on family labor and to other labor-market imperfections such as geographical dispersion and search costs that makes them unwilling or unable to leave the farm (Carter & Mesbah, 1990; Hayami, 1998). At the same time, they face a relatively high price for capital because of capital-market imperfections such as credit-rationing, transaction costs in the administration of small loans, or covariate risk (Besley, 1998); and finally, small farmers also face a high reservation demand price for land, which includes the net present value of agricultural production and its expected appreciation, plus other consumption values—such as power and prestige—which make land costly for small farmers without substantial non-covariate (i.e., outside of agriculture) amount of wealth (Binswanger, 1987; Carter & Mesbah, 1990; Henderson et al., 2015).

On the other side, large farmers face a higher shadow price for labor due to high monitoring costs of hired workers across multiple tasks and disperse rural areas, and lower opportunity costs for capital and land—the idea of two different interest rates (r^{max} and r^{min}) faced by small and large farms is elegantly formalized by Allen & Lueck (1998); and a model of differing land reservation demand prices, where the large farmers are able to out-bid the small ones, is presented by Carter & Kalfayan (1989). The fundamental aspect of the argument is that, because of these differing shadow prices, small farms commit more labor to production than large farms, while the latter treat land as a relatively abundant resource even in land-scarce economies—and substitute machines for labor even in capital-scarce, labor-abundant economies—leading to overall economic and social inefficiency (Ellis, 1993). Of course, if it weren't for factor-market imperfections, competition would bring all prices to their social optimums. Yet the existence of these factor market imperfections in developing countries is widely documented in the agricultural economics literature.

The above description of bimodal agrarian structures is clearly exemplified by the highly unequal Latin American rural sectors, where a rigid *latifundia-minifundia* social structure predominates (García, 1967: Chonchol, 1996). According to the theory, a lower shadow price for land turns the output composition of very large farms or latifundia towards land-extensive activities such as livestock pasturing—the stereotypical case of the Latin American *hacienda*, described later in this work (Reyes, 1978)—leading to a ubiquitous underutilization of land. In the meantime, labor remains unmarketably stuck in the minifundia social constellation of peasants, oftentimes in the form of disguised unemployment—i.e., peasants whose marginal product is virtually zero, or even negative (Nurkse, 1966)—and making a 'millimetric' use of the soil available to them (Vasco, 1978). All of which translates into an inefficient use of the available resources at an aggregate level—i.e., *technical inefficiency*.

Therefore, in this work we take on the empirical question of whether inequality leads to technical inefficiency, as the agricultural economics theory suggests, using a sample of Latin American rural sectors. In particular, our identification strategy relies on the historical origins of land inequality in the region and on the rigidness of the latifundia-minifundia social structure, which determine present levels of rural income inequality (Chonchol, 1996; Sokoloff & Engerman, 2000; Frankema, 2005, 2008). We also make use of a stochastic frontier model, which allows us to identify the determinants of inefficiency. We turn to these interesting methodological issues in the next section.

1.3. Methodology

To investigate the relationship between inequality and technical inefficiency empirically, we make use of stochastic frontier analysis (SFA). This technique uses output and input information of different production units to estimate a common production function with an arbitrarily chosen functional form. Following the exposition of Coelli et al. (2005), let us begin by considering the following (log-linearized) Cobb-Douglas specification:

$$ln q_i = \boldsymbol{x}_{\boldsymbol{i}}'\boldsymbol{\beta} - u_i \quad (1)$$

where q_i is the output of the *i*-th firm; $\boldsymbol{x}_{\boldsymbol{i}}$ is a $K \times 1$ vector of the logarithms of inputs; $\boldsymbol{\beta}$ is a vector of unknown parameters; and u_i is a non-negative random variable associated with technical inefficiency—i.e., the inability of the *i*-th firm to extract maximum output from a given amount of inputs. In this model, q_i is bounded from above by the non-stochastic quantity $exp(\boldsymbol{x}_{\boldsymbol{i}}'\boldsymbol{\beta})$; hence, it ignores all possible random shocks to the production process. As a response, Aigner, Lovell, & Schmidt (1977) and Meeusen & van den Broeck (1977) introduce another random variable representing statistical noise, obtaining the stochastic frontier model:

$$ln q_i = \boldsymbol{x}_{\boldsymbol{i}}'\boldsymbol{\beta} + v_i - u_i \quad (2)$$

Equation (2) differs from equation (1) in that it incorporates a symmetric random error, v_i, independent of the non-negative technical inefficiency, u_i, capturing random shocks to production and other sources of statistical noise. Together, the v_i and u_i form a composite random error, ε_i, and $v_i \sim iidN(0, \sigma_v^2)$ ^ $u_i \sim iidN^+(0, \sigma_u^2)$. In (2), the q_i is now bounded from above by the stochastic quantity $exp(x_i'\beta + v_i)$. To illustrate, imagine two firms producing homogeneous output

q_i from a single input x_i, with identical Cobb-Douglas production functions. The statistical (log-linearized) model would be: $lnq_i = \beta_0 + \beta_1 lnx_i' + v_i - u_i$. Taking exponentials:

$$q_i = exp(\beta_0 + \beta_1 lnx_i' + v_i - u_i) \qquad (3)$$

Equation (3) represents the observed levels of production, marked with a squared dot in Figure 1. The fully efficient $(u_i = 0)$—unobserved—production levels, $q_i^* = exp(\beta_0 + \beta_1 lnx_i' + v_i)$, are represented by round dots; while the deterministic part of the frontier, $exp(\beta_0 + \beta_1 lnx_i)$, is represented by the solid diminishing returns to scale curve.

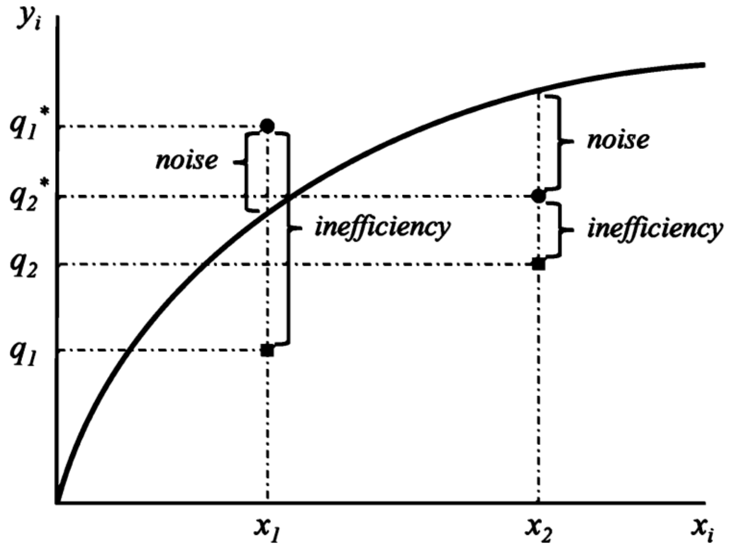


Figure 1.1. The stochastic production frontier

Adapted from Coelli et al. (2005)

Model (3) can be estimated using Maximum Likelihood (ML), with $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\lambda^2 = \sigma_u^2/\sigma_v^2$ (Sickles & Zelenyuk, 2019). While the technical inefficiency of a particular production unit (u_i) cannot be consistently estimated, Jondrow et al. (1982) suggest a consistent estimator for individual inefficiencies conditional on statistical noise, $E(u_i/\varepsilon_i)$. The resulting estimates could be further regressed on other exogenous covariates of interest to explain technical inefficiency—such as the Gini index (i.e., a measure of inequality). However, this two-stage approach is problematic: since the u_i is assumed to be uncorrelated with the $\boldsymbol{x_i}$, any other variables explaining inefficiency must be included in the first stage; otherwise, the resulting estimated parameters are biased (Greene, 2005b). To solve this issue, Battese & Coelli (1995) include appropriate explanatory variables for technical inefficiency in a single-stage approach. Their basic *panel data* model is:

$$\begin{gathered} q_{it} = exp(\boldsymbol{x_{it}\beta} + v_{it} - u_{it}) \\ u_{it} = \boldsymbol{z_{it}\delta} + w_{it} \end{gathered} \tag{4}$$

Where u_{it} is a function of its own $\boldsymbol{z_{it}}$ explanatory variables, with random error w_{it}, defined by the truncation of the normal distribution with zero mean and variance σ_u^2. This is consistent with u_{it} being a non-negative truncation (at zero) of the normal distribution with mean $\boldsymbol{z_{it}\delta}$ and variance σ_u^2 $[u_{it} \sim iidN^+(\boldsymbol{z_{it}\delta}, \sigma_u^2)]$. Again, the model is estimated using ML.

However, an important aspect of our case study is the presence of time-invariant heterogeneity across production units. Indeed, our empirical application is based on a sample of Latin American rural sectors with very disparate geographical, cultural and institutional characteristics. To capture this cross-sectional heterogeneity, Greene (2005a,b) proposes the following 'true fixed effects' (output-oriented[2]) model:

$$q_{it} = \alpha_i + \beta' \boldsymbol{x_{it}} + v_{it} - u_{it} \tag{5}$$

---

[2] In the original specification, the last term of the right-hand side of equation (5) is $-Su_{it}$, where S takes a positive (+1) or negative (−1) value depending on whether the frontier describes production or cost, respectively. In this work, we maintain an output orientation throughout.

This model retains the same distributional assumptions as before, but incorporates a country-specific, time-invariant, constant term α_i, and is estimated using a Maximum Likelihood Dummy Variable (MLDV) approach (Greene, 2005b). We utilize this model in our empirical application[3].

Nonetheless, before proceeding to our case study, let us briefly discuss some endogeneity issues. To start with, a commonly cited issue of fixed effects estimation is the so-called ‘incidental parameters problem’: with small T, many fixed effects estimators are inconsistent and subject to a small sample bias. In this regard, Greene (2005b) uses Monte Carlo simulation with $T = 5$, finding that the ‘true fixed effects’ estimator is not biased. By contrast, the variance estimators were affected substantially, causing a 25% upward bias in the efficiency effects. However, Belotti et al. (2013) point out that the MLDV approach is appropriate when $T \geq 10$; and so, we do not expect the incidental parameters problem to show up in our empirical application where we use a longer time frame ($T = 17$).

Besides the small sample bias discussed above, two basic types of endogeneity may arise in the context of stochastic frontier models (Sickles & Zelenyuk, 2019): on the one hand, Type I endogeneity occurs when the regressors are correlated with the two-sided error term; i.e., when unobserved productivity shocks (contained in the stochastic error term) affect input decisions. However, simultaneity is avoided in our case study because, at a macroeconomic level, inputs are not easily adjusted in response to random productivity shocks (e.g., total agricultural land and capital are fixed in the short-run and labor is predetermined by the existing economically active

---

[3] The authors also propose an alternative ‘true random effects’ model (Greene 2005a,b), which is rejected by a Hausman test in our empirical application.

rural population). On the other hand, Type II endogeneity occurs when the two-sided error term—the statistical noise—is correlated with the one-sided error—the inefficiency. It is here where our choice of the Latin American case study becomes instrumental, given the *exogenous* institutional and historical origins of land and rural income inequality in the region. We make an in-depth exposition of this subject in the following section.

1.4. Case Study

1.4.1. Latin American rural sectors

In this case study, we rely on the fact that land and rural income inequality are largely determined by *exogenous* institutional and historical factors. In particular, land in Latin America has traditionally been locked-up in a *rigid* bimodal latifundia-minifundia social structure: the latifundia or *haciendas* (a generic category including diverse social structures such as the cattle ranch or *hato ganadero*, the lordship [*señorial*] hacienda, or the *plantation* hacienda), on the one hand, are estates with a disproportionately large amount of land[4], and where, historically, land hasn't played a major role as a factor of production, but rather, as an element of status and political power. This lack of capitalistic rationality derives from the colonial institutions of the *Encomienda* (the lordship or

---

[4] The amount of land is, of course, a relative concept: in the Uruguayan pampas or in the Beni pastures of Bolivia, 'small' livestock operations consist of small herds in extensions of up to 500 hectares, whereas in the Valle Central of Chile or the Sabana of Bogotá, a dairy operation may not require more than 50 hectares. Hence, what is proposed as an element of characterization of the latifundia is the 'proportionality' between the magnitude of the economic operation and the amount of land (García, 1967).

personal authority over the indigenous settlements); the *Mita*, regulating the working obligations of the indigenous population; and the *estancias*, *composiciones* and *mercedes de tierras*, which added full rights over the land; effectively merging the rights over the land with the rights over the labor force. This workforce is comprised of a broad social constellation of *indios*, *colonos*, *aparceros*, *peones* and so on (different forms of peasantry, or *gleba*), isolated from the modern labor market, and immerse in a complex system of paternalistic relations (García, 1967; Reyes, 1978; Chonchol, 1996).

The counterpart to the extreme concentration of land by the latifundia, on the other hand, is the minifundia: a diversity of land tenure arrangements (*colonato*, *aparcería*, *arrendamiento*, etc.), located at the margins of the hacienda, and sustaining its *reserve army* of peasant laborers. It is fundamentally characterized by a land size incapable of providing a minimum subsistence level; its location in often steep, eroded and rain-fed hillside areas; its high level of disguised unemployment; and its level of subordination within the local structure of power. Indeed, the minifundia is not only a 'precarious' form of tenancy, but also a method of social domination within the closed social structure of the latifundia: the peasant is 'embedded' within an *asymmetric* relationship of dependence with the landlord, based on the exchange of personal services *and* political loyalty for the satisfaction of basic needs (land, food, shelter and so on), through the established paternalistic mechanisms or *clientelism* (Cendales, et al., 2022). Finally, the minifundia lacks any internal forces of change, given its tendency towards further fragmentation (due to demographic growth and the consequent succession pressures), and its sheer incapacity to generate any agricultural surpluses. Hence, it is permanently 'trapped' within the latifundia structure of power (García, 1967; Vasco, 1978; Chonchol, 1996).

Perhaps one of the most staggering facts of Latin American history is the perdurance of the latifundia-minifundia social structure, which has resisted the social, political and technological transformations of the 20th and 21st centuries. In particular, the rural reforms implemented in the continent between the 1950s and 1970s had a very limited impact in some countries due to technical and political obstacles (as in the cases of Bolivia, Brazil, Ecuador and Venezuela), or were met with violent counter-reforms (as in Colombia, Chile and Guatemala). Therefore, the processes of modernization (i.e., the adoption of green revolution technologies, the proletarianization of the labor force, the integration to the world economy, etc.) has been confined to *certain* and very *localized* agro-industrial complexes (e.g., the Chilean Valle Central, the Colombian Coffee Belt and the Brazilian soybean region of Cerrado, etc.), and contributed to exacerbate the concentration-exclusion from the land (Chonchol, 1996). Furthermore, the traditional role of land as an element of power and prestige has been reinforced with its new 'conservative' function of treasuring its continuous appreciation (Suescún, 2013). Hence, we conclude, the latifundia-minifundia social structure comprises a long-run stable equilibrium.

From the above description of Latin American rural sectors, there are at least three observations that contribute to our assertion that Type II endogeneity is not a critical issue for our SFA estimations: a) land inequality in the region is determined by very long-run processes of institutional formation. At turn, land distribution patterns determine current levels of rural income inequality[5] (Sokoloff & Engerman, 2000; Frankema, 2005; 2008), making them highly exogenous to the short-run vicissitudes of the agricultural production process itself (contained in the two-sided error term); b) the rigidity of the latifundia-minifundia social structure—in combination with the multiplicity of rural factor-market imperfections discussed in the literature review section—makes

---

[5] Here, we are implicitly acknowledging the role of factor endowments as the main determinant of income (Chenery et al., 1976; Borrero & Garza, 2019). This is particularly true in developing-country agriculture, where land is often the most important asset and where capital and land markets are imperfect or inexistent (Carter & Mesbah, 1990; Deininger & Feder, 2001).

this distribution relatively *stiff* and irresponsive to productivity shocks; and c) land accumulation patterns in the region have not always followed a capitalistic rationality, being motivated by pre-capitalistic ideologies of power and prestige.

1.4.2. Data and sources

For our SFA application, we use an (unbalanced) panel database of fifteen Latin American agricultural sectors[6] during a period of seventeen years (2005–2021). As dependent variable, we use the gross production value of agriculture (*Value*) in 2015 constant million dollars. As regressors, we include five agricultural inputs: total agricultural land (*Land*) in thousand hectares; net capital stock in agriculture (*Capital*) in 2015 constant million dollars; rural labor force (*Labor*), including all rural workers aged 15 and over; tons of fertilizers (Nitrogen, Phosphate and Potassium), pesticides and manure (*Materials*); and the weighted number of animals (*Livestock*)[7]. All the variables were retrieved from the Food and Agriculture Organization (FAO), except for the rural labor force, which was obtained from the International Labor Organization (ILO).

---

[6] Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru and Uruguay.

[7] Given the multiple species of animals, we take the approach of Hayami & Ruttan (1970) of summing up all units after assigning different weights to each animal type (e.g., 0.8 for cattle, 0.2 for pigs, 0.01 for poultry and so on).

To measure inequality—our main determinant of technical inefficiency—we use the rural income Gini concentration index[8] (*Gini*), published by the Economic Commission for Latin America (ECLAC). We add a few other determinants of inefficiency commonly used in the SFA literature: the illiteracy rate of the rural population aged 15 and older (*Illiteracy*), also from the ECLAC, as a measure of human capital (e.g., Mastromarco & Ghosh, 2009; Asefa, 2011); the simple mean of the import tariff rates applied to primary products (*Tariffs*) as a measure of exposure to international trade—higher tariffs are expected to reduce international competition and, thus, increase technical inefficiency (e.g., Frick & Sauer, 2017; Sunge & Ngepah, 2020)—retrieved from the World Bank (WB); and the proportion of total agricultural land equipped for irrigation (*Irrigation*), from FAO (e.g., Kea et al., 2016; Morais & Braga, 2018). Descriptive statistics are presented in Table 1.1.

---

[8] In a 0–100 scale, since in its original dimension (between 0 and 1), a marginal change of the Gini coefficient by 1 would not have a meaningful interpretation (Luebker, 2010).

Table 1.1. Descriptive statistics of 15 Latin American agricultural sectors (2005–2021)

Variable	Units	Mean	Std. Dev.	Min.	Max.
Frontier					
Value	constant US$	27,353	58,177	833.4	279,519
Land	hectares (ha.)	33,801	59,217	1,196	239,370
Capital	constant US$	12,092	17,957	1,371	79,458
Labor	no. of workers	3,128	3,893	92.58	16,739
Materials	tons (tn.)	1.518e+06	3.946e+06	34,704	2.385e+07
Livestock	weighted sum	3.274e+07	6.219e+07	2.128e+06	2.856e+08
Inefficiency					
Gini	index (0–100)	45.37	5.628	31.50	60.50
Illiteracy	percentage (%)	15.68	6.198	1.900	30.70
Tariffs	percentage (%)	6.926	2.753	0.36	12.89
Irrigation	percentage (%)	5.973	6.258	0.450	32.082

1.4.3. Empirical specification

As described in the methodology section, we use Greene's (2005a,b) true fixed effects model[9] (equation 5) under a log-linearized Cobb-Douglas specification[10]. Following Battese & Coelli

---

[9] As mentioned in note 3, we tested the null hypothesis that the idiosyncratic fixed effects (the α_i's) and the regressors of the main equation (i.e., the factor inputs) are uncorrelated through a Hausman test. The strong rejection of the null indicates that the 'true fixed effects' model—as opposed to the 'true random effects'—is appropriate.

[10] The Cobb-Douglas functional form has the virtue of simplicity, as the slope coefficients of the stochastic frontier can be directly interpreted as input-output elasticities. Also, a Likelihood-ratio test strongly favours

(1995) and Belotti et. al. (2013), we also add year dummies in the stochastic frontier capturing Hicks-neutral technical progress (i.e., vertical shifts of the frontier curve). The model is given by:

$$ln(Value)_{it} = \alpha_i + \beta_1 \ln{(Land)_{it}} + \beta_2 \ln{(Capital)_{it}} + \beta_3 ln(Labour)_{it} \\ + \beta_4 ln(Materials)_{it} + \beta_5 ln(Livestock)_{it} + \beta_6 Year_t + v_{it} - u_{it} \quad (6)$$

$$u_{it} = \delta_0 + \delta_1 Gini_{it} + \delta_2 Illiteracy_{it} + \delta_3 Tariffs_{it} + \delta_4 Irrigation_{it} + w_{it}$$

Before proceeding, two observations are in order: first, note that if $u_{it} = 0$, model (6) collapses to a simple log-linearized Cobb-Douglas specification; and second, recall that the fixed effects capture the cross-country time-invariant heterogeneity, which we presume to be great among Latin American agricultural sectors. Among such idiosyncratic factors are the geography, climate, political system, institutions and culture, which may vary substantially across countries (e.g., agriculture in the Andean countries is considerably different from agriculture in the pampas of the south-eastern cone). To check for the robustness of our estimates, we regress model (6) under four different specifications (models A through D), including different sets of inefficiency determinants. The results are presented in the following section.

1.5. Results

Table 1.2 displays the regression outcomes of the four SFA specifications.

---

this specification over the alternative ‘Trans-log’ functional form (χ^2_{13} = –611.88, prob>χ^2 = 1.000)—here, a negative chi-squared test statistic strongly points to not rejecting the null hypothesis.

Table 1.2. Estimation results of 15 Latin American rural sectors (2005–2021)

model VARIABLES	(A) *ln(Value)*	(B) *ln(Value)*	(C) *ln(Value)*	(D) *ln(Value)*
Countries	15	14	13	13
Observations	199	174	151	151
Prob > chi2	0.000	0.000	0.000	0.000
Log-likelihood	190.967	116.521	-141.036	-151.456
Stochastic frontier				
ln(Land)	0.046*	0.034**	0.437	0.518***
	(0.028)	(0.013)	(0.670)	(0.177)
ln(Capital)	0.050	0.067	0.150	0.235
	(0.213)	(0.244)	(0.546)	(0.554)
ln(Labor)	-0.096***	-0.090	0.338*	0.048
	(0.014)	(0.084)	(0.204)	(0.187)
ln(Materials)	0.418***	0.317	0.439	0.598*
	(0.069)	(0.208)	(0.310)	(0.310)
ln(Livestock)	0.293***	0.463***	0.428	0.415**
	(0.032)	(0.055)	(0.489)	(0.182)
Inefficiency equation				
Gini	0.239***	0.168***	0.275***	0.300***
	(0.043)	(0.057)	(0.029)	(0.026)
Illiteracy		0.241***	0.133***	0.151***
		(0.090)	(0.035)	(0.038)
Tariffs			-0.002	0.003
			(0.040)	(0.042)
Irrigation				-0.074**
				(0.034)
Ancillary parameters				
Sigma-u	1.827	0.798***	0.223	0.170
	(•)	(0.149)	(•)	(•)
Sigma-v	2.06e-08	1.24e-07	0.574***	0.637***
	(•)	(2.88e-06)	(0.074)	(0.069)
Lambda	8.86e+07	6.46e+06***	0.388	0.267
	(•)	(0.149)	(•)	(•)

All models (A through D) are based on equation (6), but with different sets of inefficiency determinants. Standard errors in parenthesis. 1%, 5%, and 10% significance levels marked ***, **, and *, respectively. Constants and year dummies omitted.

The results are compelling. Starting with the stochastic frontier component, the input-output elasticities across the four models hold the expected positive signs, except for *Labor* in the first two specifications (models A and B). Interestingly, this could be evidence of 'disguised unemployment' in Latin American rural sectors—the idea that the marginal product of labor in rural areas is close to zero or even negative, implying that a large portion of the economically active population could leave agriculture without actually decreasing agricultural output, given a fixed production technology (Nurkse, 1966). Yet, its coefficient turns positive and marginally significant in model C. As for the other inputs, *Land*, *Materials* and *Livestock* have significant positive effects (at varying levels), while *Capital* is not significant in any model. Finally, we note that the sum of all input-output elasticities is not statistically different from one in any specification, indicating constant returns to scale of the production function.

Before turning to the inefficiency equation, we checked for any potential multicollinearity issues among the inefficiency determinants and found no problems (see a correlation matrix in Table A1 in the Appendix). With regards to the Gini index—our main variable of interest—we observe in Table 1.2. that it has a *positive* and very significant effect on technical *in*efficiency throughout the four specifications. These findings are consistent with our earlier theoretical discussion of developing-country agriculture—and particularly, *bimodal* agrarian structures such as the Latin American case examined here—where differing shadow prices for land, capital and labor faced by small and large farmers leads them to use resources in ways that do not reflect their social opportunity costs, causing overall allocative inefficiency. Thus, our results support the direct inequality-inefficiency hypothesis postulated by the agricultural economics literature in the context of traditional rural sectors.

With regards to the additional inefficiency determinants, we find that *Illiteracy* is also very significant and holds the anticipated positive sign—the higher the illiteracy rate, the higher the technical inefficiency—across the three pertinent specifications. Indeed, according to the literature, human capital accumulation increases the ability to use existing resources more efficiently and facilitates the absorption of new technology (Mastromanco & Ghosh, 2009; Asefa, 2011). Similarly, *Irrigation* too is very significant and holds the expected negative sign, since better irrigation is supposed to reduce technical inefficiency by, for instance, reducing risk from weather variability (Morais & Braga, 2018). Finally, *Tariffs* is not significant, and so we cannot confirm the common assertion that higher tariffs increase technical inefficiency by reducing international competition (e.g., Frick & Sauer, 2017).

A word of caution is needed at this stage: focusing on the ancillary parameters, recall that the 'lambda' corresponds to one of the parameters of the log-likelihood function ($\lambda = \sigma_u/\sigma_v$), and its significance indicates that the inefficiency terms are indeed relevant (i.e., $u_{it} \neq 0$)[11]. Unfortunately, as we see at the bottom of Table 1.2, the regression output often returns a dot (•) in place of the standard errors of the ancillary parameters, preventing us from determining the significance of the inefficiency terms in the respective models. Hence, we must take the results of those specifications with caution. Only in model B we are able to confirm the significance of λ (at the 1% level), and so we take this as our preferred model for interpretation.

Lastly, based on model B, we calculate the technical inefficiency scores via $E(u_i/\varepsilon_i)$ using the Jondrow et al. (1982) estimator—refer back to the methodology section—and present the summary statistics of the conditional technical inefficiency effects by country in Table 1.3.

---

[11] If $\lambda \neq 0$, then the variance of the inefficiency effects is also different from zero ($\sigma_u \neq 0$), implying that the one-sided error term is significant ($u_{it} \neq 0$).

Table 1.3. Conditional inefficiency effects by country (2005–2021)

Country	Mean	Std. Dev.	Min.	Max.
Bolivia	0.253	0.337	0.00	1.463
Brazil	0.215	0.331	0.00	1.429
Chile	0.083	0.109	0.00	0.227
Colombia	0.280	0.385	0.00	1.601
Dominican Republic	0.183	0.305	0.00	1.337
Ecuador	0.193	0.350	0.00	1.497
El Salvador	0.143	0.114	0.00	0.349
Honduras	0.261	0.385	0.00	1.398
Mexico	0.196	0.108	0.00	0.292
Nicaragua	0.681	0.963	0.00	1.361
Panama	0.109	0.089	0.00	0.226
Paraguay	0.326	0.366	0.00	1.615
Peru	0.310	0.355	0.00	1.604
Uruguay	0.039	0.068	0.00	0.245
Total	0.218	0.317	0.00	1.615

We observe that, on average, Latin American agricultural sectors could produce about 21.8 percent more output with the same technology and amount of inputs (land, capital, labor, etc.) available to them, or even up to 68.1 percent more in the most inefficient case—Nicaragua. Altogether, the evidence presented in tables 1.2 and 1.3 implies that reducing inequality has a huge potential for promoting agricultural development in the region by reducing technical inefficiency. This begs the question of how to break the highly *unequal* latifundia-minifundia social structure and its traditional values of land, which would at turn alleviate the factor-market imperfections inherent to bimodal agrarian structures and bring factor prices closer to their social opportunity costs.

1.6. Conclusions

In this paper we investigate the effect of inequality on *technical inefficiency*. Indeed, the agricultural economics literature suggests that differing shadow prices for land, capital and labor faced by small and large farms in *bimodal* agrarian systems plagued with factor-market imperfections leads them to use productive resources in socially and economically inefficient ways. In particular, while very large farms treat land and capital as relatively abundant resources—even in land- and capital-scarce economies—small farms commit more labor to production, oftentimes in the form of disguised unemployment.

The above situation is clearly portrayed by the highly unequal Latin American rural sectors, where land is permanently 'locked-up' within the very large farms or latifundia—which tend towards land-extensive activities—and labor remains 'stuck' within the minifundia social constellation of peasants. Given the rigidity and the historical/institutional origins of this *latifundia-minifundia* social structure, which determine current levels of land and rural income inequality, we are able to draw causal statistical inferences in the empirical part of this work.

Specifically, we employ a sample of fifteen Latin American rural sectors over the 2005-2021 period and a stochastic frontier analysis (SFA) technique, which uses output-input information to measure the level of technical inefficiency and its determinants. Our results show that, on average, Latin American sectors could have produced 21.8 percent more output with the existing technological resources and inputs available to them. As suggested by the theory, we find a very significant *positive* effect of inequality on aggregate (i.e., macroeconomic) technical *in*efficiency. Furthermore, we identify illiteracy and irrigation as other significant factors contributing to technical inefficiency. Thus, addressing technical inefficiency should encompass not only endeavours to mitigate income inequality, but also initiatives directed at enhancing the educational levels of the rural population and the agricultural infrastructure.

Of course, there are several limitations to this study: first, although our empirical strategy confirms the main prediction of a negative inequality-efficiency relationship, the precise mechanism behind it—the mismatch between shadow prices and social opportunity costs caused by factor-market imperfections—remains unexplored. Second, the present is fundamentally a static analysis, leaving aside the dynamic implications for long-run economic *growth*. And third, although we find support of the theory within the Latin American context, testing its validity for the wider community of developing countries—with different institutional environments—remains an important task. These limitations point to plausible new directions of analysis.

Chapter 2

EFFECT OF LAND INEQUALITY ON TOTAL FACTOR PRODUCTIVITY: FARM-LEVEL EVIDENCE FROM THE PERUVIAN HIGHLANDS

2.1. Introduction

Since the inauguration of the growth accounting method by Robert Solow (1957), total factor productivity (TFP)—Solow's 'residual—has been found to be the most important source of economic growth (Brada & Bah, 2014). Its relevance is even more prominent for agriculture, since demographic pressure and limited fertile area demand ever-increasing yields (Alauddin et al., 2005). Hence, it has become crucial to understand the determinants of TFP.

Of course, many determinants have been proposed by the agricultural economics literature: human capital, research and innovation, agricultural extension, sustainable practices, and so on (e.g., Ruttan, 2002; Craig et al., 1997; Hutchins et al., 2023). However, there are other social phenomena which can affect the mix of productive resources—i.e., the technology—of an economy in a fundamental way, such as an array of institutional arrangements like the tenure or property regime, the level of democracy, trade openness, or even the level of corruption and violence (Bejarano, 1997; Gingrich & Garber, 2010; Higgins et al., 2018). Among these social phenomena, one of the aspects that has received the least attention by the scientific literature has been economic inequality—especially of *land*, the primary asset in developing country agriculture (FAO, 2002). Despite of this neglect, land inequality remains one of the most contentious topics in the political debate of underdeveloped countries, which experience both high levels of land inequality and low levels of agricultural productivity (UNCTAD, 2015; Bauluz et al., 2020). Hence, this study aims at understanding the effect of *land* inequality on agricultural TFP.

To reach this objective, we use a panel of 299 'small and medium' farms specialized in transitory or annual crops (maize, cassava, potato, etc.) from the Peruvian highland region over a four-year period (2016–2019), and implement two alternative methodologies: a dynamic panel or system GMM approach, and a control function approach. Also, we measure *land* inequality using three alternative methods: the Gini, Theil, and Kakwani indexes. The results show a significant negative association between land inequality and agricultural TFP when using the Gini and Theil indexes and no relationship when the Kakwani index was used.

In what follows, an overview of the literature linking inequality with TFP is provided (Section 2.2). In Section 2.3, the two methods used to address this study's purpose are discussed: the system GMM approach (subsection 2.3.1); and the control function approach (subsection 2.3.2.). Section 2.4 describes the data and empirical strategy used; Section 2.6 presents the results; and finally, Section 2.7 concludes.

2.2. Literature Review

In general, the economics literature links inequality with total factor productivity (TFP) through various channels: first, the Classical argument states that inequality provides the appropriate *incentives* for higher effort, thus leading to higher productivity (Landreth, 1976). Therefore, attempting to reduce inequality would distort those incentives and, as a consequence, reduce productivity. This view was advanced by the Neoclassical argument of *skill-biased technological change*, according to which non-neutral (labor-augmenting) technological change (i.e., TFP improvements) increase the demand for skilled labor—as opposed to unskilled labor—thus leading to higher wage inequality. Again, any attempt to reduce inequality would lead to lower productivity (Espoir & Ngepah, 2021). In general, then, we see that in the Classical-Neoclassical arguments, there is an 'efficiency-equity tradeoff' (Lloyd-Ellis, 2003). In connection with this purported tradeoff, another stressed argument is that of endogenous tax policy, according to which the higher the (non-economic) welfare concerns of inequality—and the poorer the median voter is—the more likely the implementation of redistributive tax policies become, risking lower levels of productivity as the efficiency-equity tradeoff suggests (Bertola, 1991; Alesina & Rodrik, 1994; Persson & Tabellini, 1994).

Other studies, by contrast, stress a negative impact of inequality on productivity through diverse channels. For example, Dusha (2019) presents a model of *corruption* where the wealthiest individuals bribe the bureaucrats for lower tax rates, which at turn act as a barrier of entry for entrepreneurship. As a result, rich individuals face a lower entry productivity threshold, while the majority of the population in the middle of the productivity distribution face a higher threshold and cannot enter into production—reducing aggregate productivity. Another tenet focuses on *human capital*. For example, Lucas (1988) elaborates a model where, because of the positive externalities of human capital, the output of each firm depends on the human capital of its own workers and on the average level of human capital in the economy. Hence, as is the case with other externalities, human capital investment in the society is suboptimal. Now then, if education exhibits constant returns to scale at the aggregate level, the existence of such externalities implies decreasing returns to scale at the micro-level; and if human capital is not perfectly tradable, heterogeneity in human capital naturally leads to lower aggregate productivity. In the same vein, we find the studies of Benabou (1996) and Acemoglu (1996), also based on human capital externalities. Finally, two more branches in this literature are identified: that based on the *political instability* channel (Gupta, 1990; Benhabib & Rustichini, 1996), and that based on *rent-seeking* behavior (Stiglitz, 2012); both also predicting a negative effect of inequality on aggregate productivity.

Now then, much of the literature on the effects of inequality on economic performance focuses on economic *growth*. This literature is based on some of the arguments stressed above and some new ones. In particular, the pioneering work in this stream of research was Kuznets' (1955) proposition of an 'inverted U' relationship between inequality and growth: positive during the early stages of development, and negative during the mature stage (an argument supported by Ahluwalia, 1976; Robinson, 1976; and Gupta & Singh, 1984). Hence, the relationship between the two depends on the level of economic development. As can be inferred, the skill-biased technological change argument is closely related to this proposition: as countries begin to develop, technological change increases the demand for skilled labor and increases inequality; but as more labor shifts towards the advanced sectors, incomes in the low-skilled jobs also begin to raise due to the shrinkage of supply of unskilled labor (Galor & Tsiddon, 1997; Helpman, 1997). On the contrary, other channels under examination, such as the role of socio-political unrest (which leads to strikes, criminality, wastage of public and private resources, discouragement of investment, and other unproductive activities) or the capital market imperfections channel—which posits that the lack of access by the poor to credit reduces high-return investment opportunities—predict a negative effect of inequality on growth (for a recent and complete review on the inequality-growth literature, refer to Mdingi & Ho, 2021). A final stream of this research has focused on asset inequality instead—particularly *land* as a proxy for wealth inequality in developing countries—as it is acknowledged to be at the bottom of every income inequality (Chenery et al., 1976; Deininger & Olinto, 2000), and to have even more skewed distributions (Brada & Bah, 2014; Lloyd-Ellis, 2003). Overall, this literature finds a negative effect of land inequality on productivity (Alesina & Rodrik, 1994; Birdsall & Londoño, 1997; Mo, 2003; Fort, 2007).

Besides the long-run growth studies, a *historical* political economy channel has also been suggested by Sokoloff & Engerman (2000), also with *land* as its central element. Specifically, the authors posit that initial land inequality conditions determine very long-run institutional quality;

which at turn, affects the very long-run economic performance of a country. Using the example of the American colonies, they explain that it was not the origin of the colonizers (as proposed by other strands of thought) which determined their long-term economic success or failure (e.g., Jamaica, unlike the other British colonies of the north, did as bad as any other Spanish colony of the Caribbean), but rather, the initial levels of land, physical, and human capital inequality—reproduced through the establishment of exclusionary institutions—which hindered long-run economic performance. On the same vein, other notable examples include the studies on the divergent paths of Canada and Argentina (Adamopoulos, 2008) and across the different regions in India (Banarjee & Iyer, 2005); all supporting the negative relationship between land inequality and very long-run economic performance.

Finally, we have the agricultural economics perspective. In this subfield, the predominant argument linking land inequality with productivity is the *inverse land size-productivity relationship* (IFSP) (Berry & Cline, 1979; Bhalla & Roy, 1988; Rada & Fuglie, 2019). In fact, in a simple model, Vollrath (2007) shows that the land inequality-productivity and land size-productivity relationships constitute a dual problem, being consistent with each other. The question is, then, why the IFSP arises in the first place. A first theoretical response can be sought in the law of diminishing returns: effectively, if the marginal product of land is decreasing with farm size, average product per hectare will also decline, producing an IFSP relationship (Ellis, 1993). A second view is that factor-market imperfections—which are typically found in developing-country rural areas (Besley, 1998; Hayami, 1998)—cause small and large farms to confront different input prices: while small farms have a relatively low cost of labor and high prices for capital and land, large farms face a higher shadow price of labor and lower opportunity costs for capital and land (Allen & Lueck, 1998). As a result, small farms commit more labor to production than large farms, and the latter treat land as a relatively abundant resource even in land-scarce economies—and substitute machines for labor even in capital-scarce, labor-abundant countries. Overall, this results in lower aggregate productivity (Ellis, 1993).

In this work, we move away from the *historical* (very long-run) and economic *growth* (long-run) approaches described above and focus instead on the effect of land inequality on agricultural *productivity* (a short-run approach). Indeed, only a handful of studies look at this particular relationship: Vollrath (2007) resorts to data of land inequality from Deininger & Squire (1998) conforming a sample of eighty countries over a period of twenty-nine years (1958–1993) and estimates the effect of land inequality on yield (output-per-hectare) using ordinary least squares (OLS), finding a significant negative effect; in Croatia, Šergo, Poropat, & Ilak (2009) look at the effect of plough land inequality on total factor productivity of wheat farms from an agricultural survey of 2003, and Silvana & Peršurić (2013) look at the effect of land inequality on TFP of vineyards from the same survey, both finding a significant negative relationship (also applying OLS regressions). We improve on this previous literature by a) implementing a TFP approach instead of partial productivity measures such as crop yield per hectare; b) moving from simple OLS regressions to production function estimations via dynamic panel and control function approaches, which account for simultaneity and attrition bias; c) using alternative measures of land inequality—Gini, Theil, and Kakwani indexes; and d) exploiting a larger and more recent dataset from the Peruvian highlands, which allows for more efficient estimations. We turn to these methodological issues in the following section.

2.3. Methodology

2.3.1. Dynamic panel data model

Let us begin by considering the following log-linearized Cobb-Douglas production function:

$$y_{it} = \boldsymbol{x}'_{\boldsymbol{it}}\boldsymbol{\beta} + \boldsymbol{w}'_{\boldsymbol{it}}\boldsymbol{\delta} + v_{it} \quad (1)$$

where y is the logarithm of output, $\boldsymbol{x}$ is a vector of log-inputs (land, capital, labor, materials, etc.), $\boldsymbol{w}$ is a vector of other observed variables (environmental, socioeconomic, etc.) determining TFP, $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are vectors of parameters to be estimated, and v_{it} is an error term. If we estimate (1) and obtain all the coefficients, TFP can be calculated as the level of observed output not attributable to factor-inputs, as:

$$ln\widehat{TFP}_{it} = y_{it} - \boldsymbol{x}'_{\boldsymbol{it}}\widehat{\boldsymbol{\beta}}$$

$$= \boldsymbol{w}'_{\boldsymbol{it}}\widehat{\boldsymbol{\delta}} + \widehat{v_{it}} \quad (2)$$

Hence, one can predict TFP from the estimated parameters[12]. Most importantly for our analysis, the estimated coefficient of the $\boldsymbol{w}$ vector can be interpreted as its average effect on the logarithm of TFP, even though it has been obtained from a regression on output. The problem is, then, to estimate (1) consistently and efficiently. Before that, one must think about the data-

---

[12] Unless in the presence of constant returns to scale, equation (2) does not satisfy the ‘proportionality’ axiom (O’Donnell, 2015). To restore proportionality, the following transformation must be applied (Ding et al., 2016):

$$ln\widehat{TFP}_{it} = y_{it} - \frac{1}{\hat{\beta}_1 + \hat{\beta}_2 + \cdots + \hat{\beta}_k}(\hat{\beta}_1 x_{1,it} + \hat{\beta}_2 x_{2,it} + \cdots + \hat{\beta}_k x_{k,it})$$

generation process: the first consideration that stands out is the fact that a powerful explanatory variable of current aggregate output levels is its past realization, i.e., $y_{i,t-1}$, since past production volumes can accurately reflect the installed capacity for production. Introducing this variable on the right-hand side of the equation, then, yields a 'dynamic panel data model'. A second consideration is that productive units (each referred to in our sample as 'unidad agropecuaria') have idiosyncratic 'fixed effects' (μ_i) associated with different tangible or intangible resources, such as soil quality or managerial capabilities. Therefore, we have now a composite error term composed of the idiosyncratic fixed effects and the random disturbance (i.e., $\varepsilon_{it} = \mu_i + v_{it}$). A fourth consideration is that some of our input-factors may be correlated with the error term, as random productivity shocks may affect input usage—i.e., *simultaneity bias*. And finally, we need to consider the nature of the data: a large *N* (over 5,000 productive units), small *T* (only 4 years) panel. Considering these five elements, equation (1) becomes:

$$y_{it} = \alpha y_{i,t-1} + \boldsymbol{x'_{it}\beta} + \boldsymbol{w'_{it}\delta} + \mu_i + v_{it} \qquad (3)$$

The next question is which estimation technique suits better the characteristics of this model: even assuming—for the moment—that all other regressors are exogenous, it is clear that simple OLS estimation would yield biased results, since the lagged value of output ($y_{i,t-1}$) is correlated with the idiosyncratic fixed effects in the error term, leading to a *dynamic panel bias* (Nickell, 1981). This problem would dissipate if *T* were large, since the impact of any one year's shock on the productive unit's fixed effect would be negligible; but again, we need to consider the short nature of our panel. To tackle this endogeneity, one approach could be to transform the data by de-meaning the variables and then applying the within-groups estimator. However, this approach doesn't eliminate the dynamic panel bias, since the lagged dependent variable is still correlated with the error term: the de-meaned lagged dependent variable becomes $y^*_{i,t-1} = y_{i,t-1} - [1/(T - 1)](y_{i2} + \cdots + y_{iT})$, while the error becomes $v^*_{it} = v_{it} - [1/(T-1)](v_{i2} + \cdots + v_{iT})$ (note that using the lagged dependent variable as regressor restricts the sample to $t = 2, \ldots, T$). The problem is that the $y_{i,t-1}$ in the new de-meaned variable is correlated with the $-[1/(T-1)](v_{i,t-1})$ term in the transformed error term (Roodman, 2009a). Instrumenting $y^*_{i,t-1}$ with lags of $y_{i,t-1}$ does not solve the problem either, as they are too embedded in v^*_{it}. Hence, another sort of transformation is needed. One such alternatives is by first-differencing the data. This gives rise to the model:

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta \boldsymbol{x}'_{it}\boldsymbol{\beta} + \Delta \boldsymbol{w}'_{it}\boldsymbol{\delta} + \Delta v_{it} \qquad (4)$$

Although this transformation gets rid of the fixed effects, the lagged dependent variable is still correlated with the error term, since the $y_{i,t-1}$ term in $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ is correlated with the $v_{i,t-1}$ term in $\Delta v_{it} = v_{it} - v_{i,t-1}$. Likewise, any predetermined variables in $\boldsymbol{x}$ that are not strictly exogenous—i.e., variables that are independent of current disturbances (v_{it}) but that can be influenced by past ones ($v_{i,t-1}$)—also become potentially endogenous. Nevertheless, longer lags of the dependent variable, either in levels ($y_{i,t-2}$) or in differences ($\Delta y_{i,t-2}$), and of the predetermined regressors ($x_{i,t-1}$ or $\Delta x_{i,t-1}$) are, this time, orthogonal to the error term (since they are correlated with $v_{i,t-2}$, not the $v_{i,t-1}$ in Δv_{it}), and so can be used as instruments. Nevertheless, this depends on the assumption that the v_{it} are not serially correlated; otherwise, the lagged variables—correlated with either contemporary or past disturbances—may correlate with future errors as well, invalidating them as instruments[13].

---

[13] A weakness of this transform, however, is that it magnifies gaps in unbalanced panels: if some y_{it} is missing, then both Δy_{it} and $\Delta y_{i,t+1}$ are missing in the transformed data. For this reason, a second common transformation is also used: ‘orthogonal deviations’, whereby the average of all future observations is subtracted from the contemporaneous one. Because it is computable for every observation except the last, it minimizes data loss. Also, because lagged observations do not enter the transformation formula, they are valid as instruments as well.

Based on the differences transformation, Anderson & Hsiao (1982) propose a specification that instruments for the lagged dependent variable with its longer lags, either in levels or in differences, and estimates it using two-stage least squares (2SLS). In their formulation, more than one lag can be included as additional instruments to improve estimation efficiency. However, this introduces a tradeoff between estimation efficiency and sample size, since a whole period is dropped with every lag. Moreover, Roodman (2009a) shows that this estimator, although consistent, performs rather poorly[14]. In general, 2SLS is efficient under homoskedasticity; but after differencing, the disturbances are no longer independent, as $\Delta v_{it} = v_{it} - v_{i,t-1}$ can be correlated with $\Delta v_{i,t-1} = v_{i,t-1} - v_{i,t-2}$, which shares the $v_{i,t-1}$ term.

This takes us to the Arellano & Bond (1991) specification, also based on the differences transform and the GMM estimator—hence, the 'difference GMM' denomination—which directly addresses this problem allowing for heteroskedasticity and correlation in the error structure[15] (Griffiths, 1993). Also, it addresses the tradeoff between efficiency and sample size, as the standard instrumental variables (IV)-type instrument of the 2SLS approach is replaced with a generalized method of moments (GMM)-style set of instruments, one for each period, substituting zeros for missing observations (Holtz-Eakin et al., 1988). Having eliminated this tradeoff, it becomes practical to include all valid lags of the untransformed variables (or levels) as instruments. For

---

[14] The author takes the LSDV and OLS point estimates as lower and upper bounds, respectively, of the credible point-estimate range. In his example of employment, this range (for the lagged employment level) is 0.733-1.045, whereas the 2SLS point-estimate is of 2.308. Moreover, for stability of the AR(1) process, this coefficient must not be greater than 1.

[15] Including a time fixed effect can help mitigate heteroskedasticity by capturing some of the systematic time-specific variations in the error structure. Time dummies may also help prevent contemporaneous correlation across individuals.

endogenous variables (such as variable inputs, thus tackling simultaneity bias), that means lags 2 and up. For predetermined but not strictly exogenous variables, that means lags 1 and longer[16].

Arellano & Bond (1991) compare the performance of difference GMM with OLS, within-groups, and Anderson-Hsiao specifications using Monte Carlo simulations, and find that difference GMM exhibits the least bias (Roodman, 2009a). However, Blundell & Bond (1998) demonstrate in other simulations that if y is close to a random walk, then difference GMM performs poorly because past levels convey little information about future changes, so untransformed lags are weak instruments for transformed variables. To increase efficiency, they develop an approach outlined in Arellano & Bover (1995), where instead of transforming the regressors to eliminate the fixed effects, it is the instruments which are transformed (differenced) to make them exogenous to the fixed effects[17]. So, while Arellano-Bond instrument differences with levels, Blundell-Bond instrument levels with differences. Indeed, for random walk-like variables, past changes may be more predictive of current levels than past levels are of current changes. In general, if w_{it} is endogenous, $\Delta w_{i,t-1}$ is available as an instrument because $\Delta w_{i,t-1} = w_{i,t-1} - w_{i,t-2}$ is not correlated with v_{it}. If w_{it} is predetermined but not strictly exogenous, the contemporaneous $\Delta w_{it} = w_{it} - w_{i,t-1}$ is available as an instrument because $E(w_{it}v_{it}) = 0$ (in either case, earlier lags are

---

[16] Because the number of instruments is quadratic in T, one can limit the number of instruments by restricting the lag ranges or by 'collapsing' the instrument set (it is important that the instrument count be less than N. Yet, in our empirical application we count with a very large number of productive units ($N > 5{,}000$), so we do not expect the number of instruments to be a problem.

[17] This conclusion is valid so long as the changes in any instrumenting variable, w_{it}, are uncorrelated with the fixed effects ($E[\Delta w_{it}\mu_i] = 0$). If this holds, $\Delta w_{i,t-1}$ is a valid instrument for the variable in levels.

available as instruments as well). Again, the validity of these statements, however, depends on the assumption that the v_{it} are not serially correlated[18].

To exploit the new moment conditions for the data in levels while retaining the original Arellano-Bond conditions for the transformed equation, Blundell & Bond (1998) designed the 'system GMM' estimator. This involved building a 'stacked' dataset with twice the observations: the untransformed and the transformed. This 'system' or augmented dataset, though, is treated as a single equation problem, since it is believed that the same linear relationship with the same coefficients applies to both types of variables. As for the instruments, the Arellano-Bond instruments for the transformed data are set to zero for the levels observations, and the Blundell-Bond transformed instruments for the levels data are set to zero for the differenced observations. A full GMM-style set of levels and differenced instruments results (excluding all the ones that result mathematically redundant). As a consequence, the system GMM estimator surpasses in efficiency the difference GMM estimator (Roodman, 2009a). In addition, it avoids the loss of data from differencing in unbalanced panels (see footnote 13) and allows for the introduction of time-invariant regressors which would otherwise disappear with only the transformed data (under the assumption that all instruments for the levels equation are orthogonal to the fixed effects (see footnote 17). We use this latter specification in this work, using the 'xtabond2' command in Stata[19].

So far, though, we have only considered two major sources of bias: dynamic panel bias and simultaneity bias, which we have tackled with our choice of the system GMM estimator. There is, however, a third source of bias that needs to be accounted for: attrition or *survival bias*. Intuitively,

---

[18] In addition, the proposition that the instrument is orthogonal to the error term requires 'stationarity' of the data generation process: for $E\left[\Delta y_{i,t-1}\varepsilon_{it}\right] = 0$ to hold, it is required that, in expectation, $|\alpha| < 1$ (Blundell & Bond, 1998).

[19] For a complete guide on the use of the 'xtabond2' command in Stata, refer to Roodman (2009).

as less productive farms exit the market and, as a consequence, drop-out of the sample, those surviving farms remaining in the sample will systematically be more productive—thus biasing our production function estimates upwards (Van Beveren, 2012). Therefore, to account for this form of bias, we follow the Olley & Pakes (1996) approach of assuming that farm i decides to stay in the market ($\chi_{it} = 1$) if its productivity (ω) exceeds a threshold of $\overline{\omega}_{it}$, which is a function of the input vector, $\boldsymbol{X_{it}}$, or to exit otherwise ($\chi_{it} = 0$); implying that at any period *t*, the probability of exiting the market is $P_{it} = P(\chi_{i,t-1} = 0)$, and thus it is a function of last year's inputs, $\boldsymbol{X_{i,t-1}}$. In practice, this entails creating a survival dummy variable (1 if the farm decides to stay in the next period, and 0 otherwise), and regressing it against the lagged inputs. The predicted value of this regression is then utilized as an additional control variable in our estimations.

2.3.2. Control function approach

As an alternative to system GMM, we use a 'control function' approach; yet again, with the main goal of tackling the input simultaneity bias problem. Once more, refer to the simple Cobb-Douglas case of equation (1), where now the error term, v_{it}, contains both factors that are completely random and unpredictable (i.e., e_{it}, a purely stochastic component), and others that, although unobservable to the econometrician, can be foreseen by the farmer (i.e., ω_{it}, a predictable productivity shock), allowing him to adjust his input mix as a response. If so, the independence assumption between inputs and the error term is violated. Thus, we may specify the Cobb-Douglas equation as:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + e_{it} \quad (5)$$

where k, l, and m are the logarithms of the fixed input (e.g., land or capital), labor, and intermediate inputs. The goal of the control function approach is, then, to find an appropriate instrument for ω_{it}. Olley & Pakes (1996) propose a two-stage procedure where *investment* is used as a proxy for the observed productivity shocks; and later, Levinsohn & Petrin (2003) propose the use of *intermediate inputs* as a proxy given the lumpiness (i.e., unresponsiveness) of investments, also in a two-stage process. However, both approaches were criticized by Ackerberg, Caves, & Frazer (2006) on the grounds that the coefficient of the labor input—also a function of the productivity shocks—may not be appropriately identified in the first stage.

In response, Wooldridge (2009) proposes a one-step approach using a GMM estimator. Its starting point is the description of the productivity process as the sum of a first-order Markov process and an 'innovation' component (φ_{it}), independent from past productivity; i.e.,

$$\omega_{it} = E\left(\omega_{it} | \omega_{i,t-1}\right) + \varphi_{it} \qquad (6)$$

Further, it is assumed that the 'state' variable (the 'fixed' input), as well as the past realizations of the 'freely variable' inputs (labor and intermediates), are independent from φ_{it} (although their contemporaneous levels are allowed to be correlated). The key aspect of this approach is that the productivity shock (ω_{it}) can be expressed as a function of the state variable and a proxy variable—e.g., investment, intermediate inputs, etc.—such that $E\left(\omega_{it} | \omega_{i,t-1}\right) \equiv g[h\left(k_{i,t-1}, m_{i,t-1}\right)]$ (Frick & Sauer, 2017). Equation (5) can thus be written as:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + g\left[h\left(k_{i,t-1}, m_{i,t-1}\right)\right] + \varphi_{it} + e_{it} \qquad (7)$$

Of course, the form of the 'control function' $g[h(.)]$ is unknown, and so it is approximated by a high-order polynomial of $k_{i,t-1}$ and $m_{i,t-1}$, with the contained variables acting as their own instruments. On its side, k_{it} acts as its own instrument too; and l_{it} and m_{it} are instrumented with their first and second lags, respectively. Estimation is done using GMM and the ‘ivregress’ command in Stata.

2.4. Data and empirical strategy

2.4.1. Data

In this work, we use four iterations (2016–2019) of the national agricultural survey (‘Encuesta Nacional Agropecuaria’, ENA) of Peru. This is a nation-wide representative sample of agricultural producers and their productive units (‘unidades agropecuarias’), dispersed across the three major geographical zones: coasts, highlands, and jungles. This dataset includes information about outputs and inputs of the agricultural production process, as well as on the socioeconomic characteristics of the producers[20]. Given the large heterogeneity of productive processes included in the sample, we focus exclusively on ‘small and medium’ transitive crop producers in the Peruvian highlands, seeking for a higher homogeneity of the production function. This restricts the analysis to a set of 299 productive units that repeat at least twice successively during the four-year period of analysis (for a total of 722 observations), yielding an unbalanced panel with the desired large *N*, small *T* structure of the system GMM approach.

To identify the set of farms, we took the approach followed in the literature (e.g., Hadley, 2006; Skevas & Martinez-Palomares, 2023) of selecting those productive units whose transitory

---

[20] The reader may refer to the ENA’s technical specification sheet for further details of how the survey is conducted (INEI, 2019).

crop output exceeded two-thirds of the total farm output (including both crops and livestock production). Again, by selecting a sample of farms specialized in transitory crops, the homogeneity of the production function is preserved. Effectively, while choosing farms specialized in a single crop would have seemed ideal from a homogeneity perspective, mixed-crop technologies seem to be the norm in traditional Peruvian-highland agriculture, as revealed by the INEI sample: the typical farmer from the Peruvian highlands produces roughly four different crops—and even up to twenty-six—with a large proportion of them (about 30%) in a disperse, garden ('vergel'), or associated fashion. In fact, according to the literature, Andean traditional agriculture is fundamentally characterized by complex intercropping systems (the *milpa*, *conuco*, etc.), which from an agronomic and economic perspective, constitute a single technological arrangement[21]. Moreover, we have narrowed the selection to 'small and medium' productive units—the largest farm has roughly 40 hectares—all of which are household operated (family members are the main source of labor).

Output is then defined as the total crop output, including the portions of output destined for the market and for self-consumption of the household, and for the production of seed, either for the market or for use as self-input. Given the variety of crops and measures (kilos, quintales, sacos, etc.), we use the sum of their monetary value ('Soles') to construct a single output variable, y. As

---

[21] For example, the customary *milpa* is the association of three key nutritional elements: maize, which provides the carbohydrates; beans, rich in proteins and oils; and calabashes, contributing with vitamins and cellulose. Altogether, they conform a single technological arrangement: the maize provides a support structure for the beans, while the abundant foliage of the calabashes protects the soil from overheating and from the impact of the rain. Other combinations, such as the *conuco*, include a wider diversity of crops; and overall, respond to carefully planned diversification strategies to achieve higher nutrition diets and risk-mitigation (Chonchol, 1996).

inputs—the $\boldsymbol{x}$ vector in the untransformed equation (1)—we include: total *land*, measured in hectares; *labor*, measured as the total number of permanent, transitory, and family workers (aged 5 years and older); and *materials*, which include a wide range of variable costs, in monetary value: seed, manure, fertilizer, pesticides, water, equipment rent and maintenance, fuel, and others. Although the survey does not collect data on the stock of capital inputs (e.g., buildings, machinery, and equipment), some capital proxies such as equipment rent and maintenance and others (e.g., yoke of oxen, etc.) are available. However, since most observations for these variables are zero in our sample, we take the approach of Skevas & Martinez-Palomares (2023) of including them in the *materials* variable. All monetary variables were deflated using the Consumer Price Index, and all input variables are expressed in logarithms, according to the specification of our models.

In the $\boldsymbol{w}$ vector of variables affecting TFP, and in line with our main research objective, we begin by including three different measures of land inequality, of which the first two are the commonly used *Gini* and *Theil* coefficients. The first of them is constructed by arranging the sample of farmers in ascendant order of individual land endowments and tracing a curve of the cumulative percentage of endowed land against the cumulative percentage of farmers (known as the 'Lorenz curve'). This curve is then compared to a 45-degree angle line of 'perfect equality', where each percentile of the farmers would have an equal share of land. The Gini coefficient is, then, calculated as the ratio of the area between the Lorenz curve and perfect equality line to the total area under the perfect equality line (refer to Luebker, 2010, for a detailed explanation), giving rise to an index between 0 and 1 (with values closer to 1 indicating greater inequality). However, we multiply this index by 100 for interpretation purposes, since a marginal change of the Gini coefficient by 1 unit would not have a meaningful interpretation (ibid.). Lastly, the index is calculated at the 'departamento' level every year and then applied to each farm individually. Therefore, every farm within each departamento (and same year) has the same Gini index. Similarly, the Theil index also measures the extent to which individual farmers within the sample differ in their land endowment levels, although its construction—based on the concept of 'entropy'—is much less intuitive (refer to Conceição & Ferreira, 2000, for a complete discussion of this index). Bounded between 0 and infinity, a higher value of the Theil index also indicates a greater disparity of the distribution of land; and as with the Gini coefficient, we multiply this index by 100 and calculate it at the departamento level.

The third measure of inequality used in this work is the *Kakwani* index (Wang et al., 2022), which we use to measure inequality at the household or farm level based on the concept of 'relative deprivation'. Specifically, this index compares the land endowment level of each productive unit within the sample against that of the remaining farms in the sample: the higher the index for a particular unit, the higher its relative land deprivation, or the more unfavorable its social position when compared to others in terms of land endowments. Following our earlier discussions, in this study we focus on land inequality (or land relative deprivation) as the main factor determining social inequities of the Peruvian rural sector.

The calculation of the Kakwani Index proceeds as follows: first, the $N = 1, 2, \ldots, n$ farms in the sample are rearranged according to their land endowments, Z, in ascending order, i.e., $z_1 \leq z_2 \leq \cdots \leq z_n$. Then, the relative land deprivation for the *i*'th productive unit is:

$$Kakwani_i = \frac{1}{N\mu_z}\sum_{j=i+1}^{n}(z_j - z_i)$$

Or, equivalently,

$$Kakwani_i = \gamma_{z_i}^{+}\left[\frac{(\mu_{z_i}^{+} - z_i)}{\mu_z}\right] \quad (8)$$

where μ_z is the average cultivated land endowment of the sample, $\mu_{z_i}^{+}$ is the average cultivated land endowment of those farms with more land than z_i, and $\gamma_{z_i}^{+}$ is the proportion of farms with more land than farm i (i.e., M/N, where $M = j, \ldots, n$). This index is bounded between zero and one. The higher the proportion of farms with more land than farm *i*, or the larger their average land compared to farm *i*'s endowment, the closer to 1 the index will be, and the worse the relative situation of farm *i* compared to the rest of the productive units in the sample. Conversely, the closer to zero, the lower the relative deprivation of the farm, and the more favorable its situation compared to the rest of productive units. For interpretation purposes, however, we also multiply this index by 100. In line with the historical, agricultural economics, and several other approaches of our literature review, we expect the coefficient of this index—as well as those of the *Gini* and *Theil* indexes—to be negative: the lower the average relative land deprivation—or level of inequality—the higher the average productivity.

Apart from inequality, there are other socio-economic variables that may affect TFP. The choice of these variables is based on previous studies examining the determinants of farm-level TFP (e.g., Vollrath, 2007; Silvana & Peršurić, 2013) and data availability. As a result the following socio-economic variables were identified as possible TFP predictors: *age* and *sex* of the main operator (1 if male, 0 if female); the operator's level of education, captured by two dummy variables—*loweduc* and *higheduc*[22]; whether the operator had his/her primary *residence* on the farm (1 if yes, 0 otherwise); and the *distance* to the district capital from the nearest plot, in hours of travel. Also, we include a set of three continuous variables measuring the percentage of land under each type of tenancy arrangement: owner or *propietario*, rentee or *arrendatario*, and *traditional* tenures, which include communal tenants ('comuneros'), share tenants ('posesionario'), and others.

---

[22] The original ENA survey measures education in ten discrete levels. For parsimony of the model, we regroup them into these two categories, where *loweduc* includes no level, initial education, incomplete primary education, complete primary education, and incomplete secondary education; and *higheduc* includes complete secondary education, incomplete superior non-university education, incomplete superior university education, complete superior non-university education, and complete superior university education.

These are calculated by dividing the number of hectares of land under each form of tenancy over the total number of hectares of the productive unit, times 100. As opposed to the $\boldsymbol{x}$ vector, none of the variables contained in $\boldsymbol{w}$ are expressed in logarithms.

There is, apart from the $\boldsymbol{x}$ and $\boldsymbol{w}$ vectors, one more variable to be included in the regression analysis. As it is discussed earlier, the most productive farms are more likely to remain in the sample, thus creating a selection bias problem. Therefore, a survival dummy variable (1 if the farm decides to stay in the next period, 0 if it decides to exit) was created and regressed against the lagged inputs (*land*, *labor*, and *materials*). The predicted value of this regression, *attrition*, reflects the probability of survival of any given farm in the market (refer back to section 2.3.1), and was used as an additional control variable in the estimations. The full variable list and descriptive statistics are presented in Table 2.1.

Table 2.1. Descriptive statistics

VARIABLES	Units	Mean	Std. Dev.	Min.	Max.
Output/inputs					
y	constant Soles ($)	1,446	8,084	1.141	211,186
land	hectares (ha.)	1.141	2.303	0.006	39.75
labor	workers	5.057	4.441	1	26
materials	constant Soles ($)	519.6	1,144	1.113	21,800
Socioeconomic variables					
Gini	index (0–100)	58.66	12.41	23.68	81.84
Theil	index (0–∞)	73.16	38.05	11.68	182.3
Kakwani	index (0–100)	75.90	18.76	5.489	99.76
age	years	55.91	16.39	21	93
sex	dummy (base: female)	0.670	0.470	0	1
higheduc	dummy (base: loweduc)	0.197	0.398	0	1
propietario	percentage (%)	67.33	43.32	0	100
arrendatario	percentage (%)	4.468	16.50	0	100
residence	dummy (base: no)	0.378	0.485	0	1
distance	hours	1.364	1.551	0	13
Other variables					
attrition	Probability (0–1)	0.587	0.019	0.507	0.800

2.4.2. Empirical strategy

2.4.2.1. Dynamic panel data model

For estimating our dynamic panel data model using system GMM, certain provisions must be made: following Roodman (2009a), we treat the lagged dependent variable ($L.y$) as a predetermined but not strictly exogenous variable; while all the inputs in the $\boldsymbol{x_{it}}$ vector are treated as endogenous variables. As it is discussed earlier, the reason $\boldsymbol{x_{it}}$ are assumed to be endogenous is that farmers may adjust their input decisions in accordance with random productivity shocks, contained in the error term (i.e., simultaneity bias). Regarding our inequality measures, they receive distinct treatments: whereas both the *Gini* and the *Theil* coefficients are treated as exogenous, the *Kakwani* index is treated as endogenous. To see why, recall that the first two indexes are calculated from the land endowments of every farm within each departamento, and consequently, the incidence of any single farm's productivity shock on the aggregate level of inequality is negligible—unless, of course, the shock is correlated across a significant number of farms[23]. By contrast, the Kakwani index relies heavily on each farm's own land information—the x_i in the numerator of equation (8)—and so it is more sensitive to productivity shocks affecting individual farms.

At turn, all but one of the $\boldsymbol{w_{it}}$ variables (*age*, *sex*, *higheduc*, etc.) are treated as strictly exogenous variables. In particular, the different forms of tenure in Peru, whether formal (*propietario*) or informal (*traditional*) are determined by very long-run processes of institutional formation, as opposed to the short-run vicissitudes of the production process (García, 1967;

---

[23] As will be seen shortly, we also include a time fixed effects variable to mitigate the possibility of contemporaneous correlation across individuals.

Zegarra, 1999). The only variable not treated as strictly exogenous but as endogenous is *arrendatario*, since the decision of renting-in land may be correlated with random productivity shocks (e.g., a "good season"). *loweduc* and *traditional* are treated as the base categories for education and tenancy arrangement, respectively.

Besides the variables contained in the $\boldsymbol{x_{it}}$ and $\boldsymbol{w_{it}}$ vectors, we also include our *attrition* variable to account for the possibility of survival bias; the idiosyncratic fixed effects—the μ_i in equation (3); and time fixed effects (t_t), to account for nation-wide exogenous phenomena that affect all productive units the same (such as a major policy change or a major climatologic event, like the El Niño phenomenon). As suggested by Roodman (2009a), this also allows to mitigate heteroskedasticity and to prevent contemporaneous correlation across individuals. The final dynamic panel data model to be estimated looks as follows:

$$\begin{aligned} y_{it} = \alpha y_{i,t-1} &+ \beta_1 ln(land)_{it} + \beta_2 ln(labor)_{it} + \beta_3 ln(materials)_{it} + \beta_4 Inequality_{it} \\ &+ \beta_5 age_{it} + \beta_6 sex_{it} + \beta_7 higheduc_{it} + \beta_8 propietario_{it} + \beta_9 arrendatario_{it} \\ &+ \beta_{10} residence_{it} + \beta_{11} distance_{it} + \beta_{12} attrition_{it} + \mu_i + t_t + v_{it} \quad (9) \end{aligned}$$

where *Inequality* stands for either of our three land inequality measures (*Gini*, *Theil*, and *Kakwani* indexes), and all other variables are as previously defined.

2.4.2.2. Control function approach

To estimate the control function model, certain provisions need to be made: *land* is specified as the 'state' variable (due to the absence of a capital variable in the data), labor as a 'freely variable' input, and *materials* as the 'proxy' variable. Thus, the productivity shock (ω_{it}) foreseen by the farm can be instrumented with the control function $g\{h[ln(land)_{i,t-1}, ln(materials)_{i,t-1}]\}$. As explained earlier in the methodological section, *land* acts as its own instrument, while *labor* and *materials* are instrumented with their first and second lags, respectively. The model to be estimated is then:

$$y_{it} = \beta_0 + \beta_1 ln(land)_{it} + \beta_2 ln(labor)_{it} + \beta_3 ln(materials)_{it} + \beta_4 Inequality_{it} + \beta_5 age_{it} + \beta_6 sex_{it} + \beta_7 higheduc_{it} + \beta_8 propietario_{it} + \beta_9 arrendatario_{it} + \beta_{10} residence_{it} + \beta_{11} distance_{it} + \beta_{12} attrition_{it} + g\{h[ln(land)_{i,t-1}, ln(materials)_{i,t-1}]\} + \varphi_{it} + t_t + e_{it} \tag{10}$$

Once again, because the form of the control function $g[h(.)]$ is unknown, we approximate it by a third-order polynomial of *land* and *materials*—following the practice of Frick & Sauer (2017). We have also introduced time fixed effects (t_t); but this time, the inclusion of idiosyncratic fixed effects is not permitted by the technique (we shall discuss the implications later). Again, we use GMM and the Stata 'ivregress' command for estimation.

2.4.2.3. Robustness checks

To check for the robustness of our estimates, several specifications are run: for comparison, we begin by estimating a simple fixed-effects (FE) model (model A), estimated with Ordinary Least Squares and the within-transformation. Essentially, it is equivalent to equation (3), but without the lagged output variable as dependent variable ($\alpha = 0$). Of course, this model is not subject to dynamic panel bias; yet, it does not account for simultaneity nor survival bias. We then re-estimate this model with our *attrition* variable to address the latter issue (model B). Following these two reference models, we turn to system GMM estimation to account for the endogeneity of inputs, also without and with the *attrition* control variable (models C and D, respectively). Finally, we employ the Wooldridge (2009) control function approach, including the attrition variable altogether (model E). We repeat these five specifications with each of our three inequality measures—the *Gini*, *Theil*, and *Kakwani* indexes. The results are presented in Tables 2.2 through 2.4.

2.5. Results

Before interpreting our results, we must conduct the rutinary post-estimation diagnosis, particularly in the case of the system GMM estimations (models C and D). Indeed, recall from the methodology section that some lags are invalid as instruments when the idiosyncratic error terms are serially correlated—for example, when $y_{i,t-2}$ is serially correlated to the $v_{i,t-1}$ in the error term in differences. In such a situation, the third lag of y must be used as an instrument; unless second-order serial correlation was found, in which case longer lags of y would be necessary. Hence, to test for autocorrelation, the Arellano-Bond test is applied to the residuals in differences. Because Δv_{it} is mathematically related to $\Delta v_{i,t-1}$ via the shared $v_{i,t-1}$ term, negative first-order serial correlation is expected in differences—as reflected by the p-value of the AR(1) tests of our models (not reported). Thus, to check for first-order serial correlation in levels, we must look for second-order correlation in differences, since this would detect correlation between the $v_{i,t-1}$ in Δv_{it}, and the $v_{i,t-2}$ in $\Delta v_{i,t-2}$ (Roodman, 2009a).

Unfortunately, since we only have three years in our sample (one year got lost because of the lagged dependent variable) and the AR(2) Arellano-Bond test requires at least four time periods to be computed, we are unable to perform this test. In fact, this issue has not been properly addressed in the previous literature, including Roodman (2009a, 2009b). To this regard, however, Mulusew & Mingyong (2023)—who face the same situation—note that large-*N*, small-*T* data is far more susceptible to heteroskedasticity than to serial correlation and suggest the use of robust standard errors to mitigate this issue, as we have done in all specifications. Finally, in addition to the Arellano-Bond test, it is standard to check for the joint validity of the instruments using the Hansen test of over-identifying restrictions. To this regard, all our system GMM estimations pass the test, failing to reject the null of overall validity of the instruments.

Finally, we checked for any potential multicollinearity issues among the productivity determinants, and found no problems (see a correlation matrix in Table A2 in the Appendix). The regression outputs are presented in Tables 2.2–2.4 below.

Table 2.2. Estimation output using the Gini index

	(A)	(B)	(C)	(D)	(E)
	FE	FE-attr.	GMM	GMM-attr.	Wooldr.
L.y			0.107	0.108	
			(0.111)	(0.113)	
materials	0.662***	0.654***	0.612***	0.634***	0.955
	(0.057)	(0.087)	(0.222)	(0.242)	(0.736)
labor	0.034	0.085	0.258	0.223	-0.317
	(0.069)	(0.117)	(0.227)	(0.223)	(1.361)
land	0.132**	0.121	0.073	0.068	0.051
	(0.055)	(0.106)	(0.103)	(0.114)	(0.303)
L.land					0.220
					(0.226)
$L.land^2$					0.131
					(0.207)
$L.land^3$					0.015
					(0.050)
L.materials					-0.389
					(1.463)
$L.materials^2$					0.030
					(0.291)
$L.materials^3$					-0.000
					(0.020)
Gini	-0.005	-0.013*	-0.011**	-0.011**	-0.013
	(0.004)	(0.008)	(0.005)	(0.005)	(0.015)
age	0.008	0.005	-0.005	-0.006	-0.015
	(0.006)	(0.010)	(0.005)	(0.004)	(0.022)
sex	0.061	0.755*	0.102	0.101	0.281
	(0.294)	(0.428)	(0.120)	(0.121)	(0.193)
higheduc	0.092	0.172	0.043	0.044	0.107
	(0.160)	(0.321)	(0.153)	(0.151)	(0.428)
propietario	0.002*	0.004**	0.002	0.002	0.003
	(0.001)	(0.002)	(0.002)	(0.001)	(0.004)
arrendatario	0.001	0.006	0.015	0.014	-0.004
	(0.003)	(0.006)	(0.013)	(0.013)	(0.044)
residence	-0.035	0.005	0.010	0.002	-0.110
	(0.115)	(0.221)	(0.098)	(0.097)	(0.199)
distance	0.057	0.093	0.037	0.035	0.116
	(0.045)	(0.066)	(0.043)	(0.043)	(0.099)
attrition		0.361		1.508	-18.25
		(4.454)		(3.091)	(47.51)
year dummies	yes	yes	yes	yes	yes
Prob > F / χ^2	0.000	0.000	0.000	0.000	0.000
R-sq. (overall)	0.541	0.515	-	-	0.460
Hansen: Pr > χ^2	-	-	0.723	0.720	-
No. of farms	299	299	299	299	92
No. of obs.	720	422	422	422	123

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2.3. Estimation output using the Theil index

	(A) FE	(B) FE-attr.	(C) GMM	(D) GMM-attr.	(E) Wooldr.
L.y			0.109	0.110	
			(0.111)	(0.113)	
materials	0.660***	0.642***	0.609***	0.630***	0.936
	(0.057)	(0.088)	(0.218)	(0.237)	(0.705)
labor	0.034	0.078	0.254	0.219	-0.235
	(0.069)	(0.117)	(0.228)	(0.224)	(1.394)
land	0.135**	0.117	0.077	0.072	0.039
	(0.055)	(0.107)	(0.101)	(0.112)	(0.305)
L.land					0.207
					(0.226)
$L.land^2$					0.119
					(0.210)
$L.land^3$					0.013
					(0.049)
L.materials					-0.295
					(1.500)
$L.materials^2$					0.015
					(0.294)
$L.materials^3$					0.001
					(0.020)
Theil	-0.001	-0.006*	-0.003*	-0.003*	-0.004
	(0.001)	(0.003)	(0.002)	(0.002)	(0.006)
age	0.008	0.006	-0.005	-0.006	-0.013
	(0.006)	(0.010)	(0.005)	(0.004)	(0.022)
sex	0.066	0.727*	0.104	0.103	0.280
	(0.292)	(0.405)	(0.119)	(0.120)	(0.186)
higheduc	0.096	0.197	0.043	0.043	0.117
	(0.159)	(0.319)	(0.156)	(0.154)	(0.427)
propietario	0.002*	0.004**	0.003*	0.002	0.003
	(0.001)	(0.002)	(0.002)	(0.001)	(0.004)
arrendatario	0.001	0.007	0.016	0.015	-0.004
	(0.003)	(0.006)	(0.013)	(0.013)	(0.041)
residence	-0.033	0.024	0.010	0.002	-0.104
	(0.115)	(0.222)	(0.099)	(0.098)	(0.196)
distance	0.056	0.106	0.036	0.034	0.113
	(0.045)	(0.065)	(0.043)	(0.042)	(0.095)
attrition		0.807		1.479	-15.50
		(4.316)		(3.076)	(48.74)
year dummies	yes	yes	yes	yes	yes
Prob > F / χ^2	0.000	0.000	0.000	0.000	0.000
R-sq. (overall)	0.539	0.501	-	-	0.488
Hansen: Pr > χ^2	-	-	0.727	0.721	-
No. of farms	299	299	299	299	92
No. of obs.	720	422	422	422	123

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2.4. Estimation output using the Kakwani index

	(A) FE	(B) FE-attr.	(C) GMM	(D) GMM-attr.	(E) Wooldr.
L.y			0.132	0.132	
			(0.116)	(0.118)	
materials	0.664***	0.644***	0.659***	0.674***	0.883
	(0.057)	(0.090)	(0.225)	(0.245)	(2.462)
labor	0.033	0.053	0.249	0.221	-0.371
	(0.069)	(0.123)	(0.254)	(0.251)	(2.782)
land	0.154	0.091	0.004	0.015	0.230
	(0.136)	(0.213)	(0.183)	(0.183)	(4.606)
L.land					0.239
					(1.536)
L.land^2					0.132
					(0.971)
L.land^3					0.015
					(0.106)
L.materials					-0.473
					(3.894)
L.materials^2					0.043
					(0.584)
L.materials^3					-0.001
					(0.041)
Kakwani	0.002	-0.003	-0.004	-0.003	0.009
	(0.009)	(0.015)	(0.009)	(0.009)	(0.368)
age	0.008	0.007	-0.005	-0.005	-0.013
	(0.006)	(0.010)	(0.005)	(0.004)	(0.027)
sex	0.066	0.460	0.094	0.093	0.254
	(0.276)	(0.433)	(0.115)	(0.116)	(0.407)
higheduc	0.092	0.171	-0.023	-0.022	0.057
	(0.159)	(0.315)	(0.153)	(0.151)	(0.441)
propietario	0.002	0.004**	0.002	0.002	0.003
	(0.001)	(0.002)	(0.002)	(0.002)	(0.010)
arrendatario	0.001	0.006	0.011	0.010	0.009
	(0.003)	(0.006)	(0.013)	(0.013)	(0.054)
residence	-0.038	-0.057	-0.032	-0.039	-0.141
	(0.115)	(0.209)	(0.099)	(0.099)	(0.237)
distance	0.057	0.086	0.034	0.032	0.096
	(0.045)	(0.077)	(0.043)	(0.043)	(0.227)
attrition		-0.561		1.175	-20.35
		(4.718)		(2.989)	(102.6)
year dummies	yes	yes	yes	yes	yes
Prob > F / χ^2	0.000	0.000	0.000	0.000	0.000
R-sq. (overall)	0.535	0.521	-	-	0.466
Hansen: Pr > χ^2	-	-	0.777	0.776	-
No. of farms	299	299	299	299	92
No. of obs.	722	423	423	423	123

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Tables 2.2–2.4 present the results of estimating models A-E for each of the three inequality indexes examined in this study. Regarding the contribution of inputs to output, *materials* has the largest input-output elasticity (roughly in the range of 0.61–0.67) and is the only statistically significant input across the four FE and system GMM regressions (models A through D), highlighting the importance of intermediary inputs (seed, manure, fertilizer, pesticides, etc.) in Peruvian agriculture. Also, the sum of the input-output elasticities of the three categories of inputs is less than one (as per a simple *t*-test), indicating decreasing returns to scale (DRS). According to the agricultural economics view (refer back to the literature review section), this would be consistent with the idea of an inverse farm-size productivity relationship—as a declining marginal product implies a decreasing average product with farm size—giving rise to an inverse inequality-productivity relationship (Vollrath, 2007).

With regards to the *Gini* and *Theil* indexes (Tables 2.2–2.3), the following pattern is observed: moving from left to right on either table, starting with model A, which does not account for any form of bias, both inequality indexes have the expected negative sign, although they are not significant. As soon as survival bias is accounted for via our *attrition* variable (model B), both indexes become marginally significant (i.e., at the 10% level); and when both dynamic panel bias and simultaneity are accounted for in our system GMM regressions (models C and D), their level of significance increases to 5%. Nonetheless, neither index is significant in the control function approach (model E)—and in fact, no variable is. Perhaps, this is due to the absence of fixed effects, which we presume to be relevant in agricultural environments; or to the dramatic loss of data points (only 123, compared to 422 and 720 of the previous models) which results from lagging the data (the *materials* variable) twice. Focusing, thus, on the former regressions, we find evidence of a 1.1 to 1.3 percent decrease in total factor productivity per each percentage point increase in the land

Gini index; and of a 0.3–0.6 percent decrease per each unit increase in the land Theil index[24]. As discussed earlier in the literature review, this negative relationship has many plausible explanations (corruption, the political economy channel, etc.). Although we are unable to determine the precise mechanism behind this relationship, the finding of DRS is at least consistent with the agricultural economics approach, according to which an inverse farm size-productivity relationship translates into a negative inequality-productivity relationship.

With respect to the Kakwani index regressions (Table 2.4), we note that *materials* continues to be the predominant input with a 1% level of significance across the first four models. Unlike the previous cases, though, the Kakwani index is not significant in any specification, although it holds the expected negative sign. Again, the Wooldridge (2009) control function approach does not give 'satisfactory' results, for the reasons discussed in the previous paragraph (the absence of fixed effects and the loss of information). Thus, we do not find evidence of a negative effect of land inequality on TFP when measuring inequality as relative deprivation.

As for the remaining socioeconomic variables, we see that most of them are statistically insignificant, with only two exceptions: on one hand, we find a marginally significant positive effect of being male (*sex*) on TFP, which is in line with findings from other studies—since, it is believed, women are bound by gender roles and imbalances of power that may not allow them to invest or barter as effectively as men (e.g., Zegarra et al., 2008; Navarro-Castañeda et al., 2021). On the other hand, *propietario* occasionally has a significant or marginally significant positive effect on productivity, as is widely reported by the property rights literature arguing that individual

---

[24] Whereas the Gini is bounded between 0 and 100, the Theil index goes from 0 to infinity, and so a one-unit increase of the Theil index does not precisely correspond to a one-percentage point increase.

property rights allow the internalization of costs and benefits, facilitate the access to credit, and promote investment (Demsetz, 1967; De Soto, 2000).

As a last step, log-TFP estimates were generated using equation (2)—in its modified form of footnote 12, given our finding of DRS—and the GMM specification of model D, Table 2.2. Average log-TFP at the beginning of the period, in 2016, was estimated at 0.95; while in 2019, it was estimated at 0.98, for an approximate 3.5 percent growth over the four-year period. Rather than solely assessing mean log-TFP values over time, a two-sample Kolmogorov-Smirnov (K-S) test is employed to compare the distributions of log-TFP at the beginning and conclusion of our study period. Figure 1 shows the cumulative distributions of log-TFP for both 2016 and 2019. The K-S test did not reject the null hypothesis of equality for the log-TFP distributions of 2016 and 2019 (D = 0.0577, p-value: 0.979). Therefore, there is no statistical evidence to support the claim that log-TFP was higher in 2019 compared to the beginning of the study period.

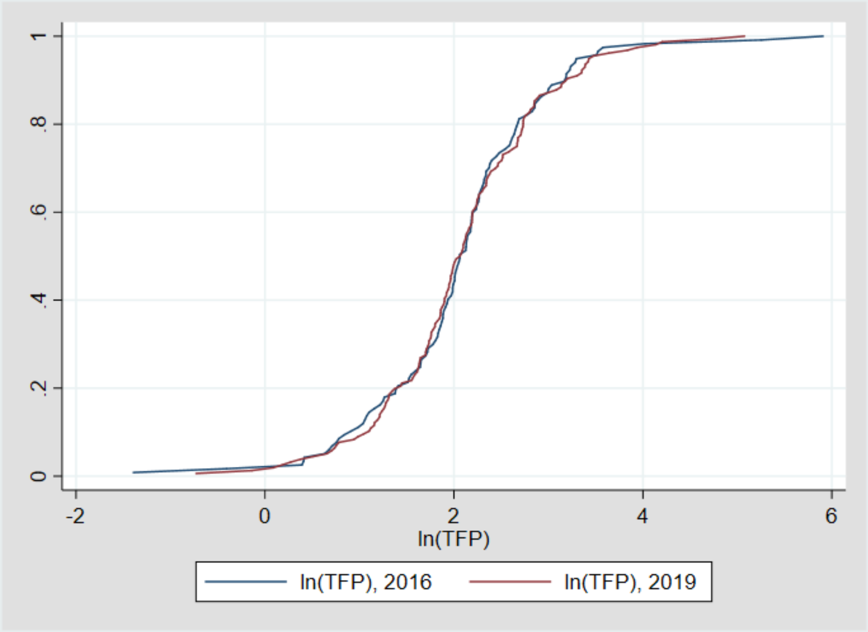


Figure 2.1. Average log-TFP cumulative distribution functions (2016 and 2019)

2.6. Conclusions

This study examines the relationship between land inequality and total factor productivity (TFP) using a sample of 299 small and medium farms from the Peruvian highlands spanning the period from 2016 to 2019. Employing models that address input endogeneity and selection bias, three different land inequality indexes are evaluated: the Gini and the Theil coefficients, and the Kakwani relative deprivation index. The findings reveal that land inequality, as measured by the Gini and Theil coefficients, has a negative impact on TFP. However, the Kakwani index does not exhibit significance in any of the estimated models, despite displaying the expected negative sign. Moreover, the study concludes that there was no notable change in productivity among the sampled farms over the period of analysis.

Overall, these findings highlight the necessity for policy interventions or support mechanisms to increase productivity levels, which have remained stagnant among the sampled farms. Such lack of productivity growth could potentially result in diminished profits, reduced competitiveness, and ultimately threaten the sustainability of agricultural production over time. Given that land inequality emerges as a significant factor contributing to this productivity decline, efforts should be directed towards reducing land inequality in the study region. These efforts may also involve initiatives aimed at securing land property rights, as this factor appears to contribute to higher productivity among the sampled farms.

Future research could explore the mechanisms underlying the inverse relationship between inequality and productivity identified in this study. This may involve investigating whether unequal land access limits smallholder farmers' adoption of productivity-enhancing practices and expansion of operations, and whether larger landholders underutilize their land or prioritize less productive uses. Additionally, it would be valuable to examine if unequal land ownership contributes to disparities in access to credit for adopting productivity-enhancing inputs or technologies, unequal market power, and access to extension services and knowledge. Understanding how these factors interact could inform more targeted interventions to reduce inequality and enhance the productivity of the sample farms. Finally, while the results obtained are valid within the specific context of Peruvian Andean agriculture, a relevant question arises regarding their applicability to other regions of Peru, such as the more export-oriented Costa region, or to agricultural sectors with very different features around the world. These areas represent promising avenues for further inquiry.

Chapter 3

THE UNEXPLORED LINK BETWEEN LAND INEQUALITY AND THE ADOPTION AND INTENSITY OF USE OF CLIMATE SMART AGRICULTURAL PRACTICES: EVIDENCE FROM A SAMPLE OF PERUVIAN FARMERS

3.1. Introduction

Human-induced greenhouse gas emissions contribute to global warming and increase the probability of weather extremes (e.g., floods, droughts, etc.) (Mirza, 2003). These phenomena are especially pervasive for agriculture, given its reliance on weather and natural resources, leading to crop failures, yield losses, soil degradation, and so on (Mizik, 2021). More importantly, these consequences affect disproportionally the world's poor, whose livelihoods rely heavily on agriculture (Bogdanski, 2012).

As a direct response to these threats, the concept of 'climate smart agriculture' (CSA) has emerged, defined as "agriculture that sustainably increases productivity, enhances resilience, reduces greenhouse gases, and enhances achievement of national food security and development goals" (FAO, 2010). According to this definition, the three pillars of CSA are: a) increasing agricultural *productivity*; b) increasing resilience, or *adaptation*; and c) reducing greenhouse gas emissions, or *mitigation* (Mizik, 2021). Thus, broadly speaking, any agricultural practice that contributes to one or more of these pillars can be categorized as CSA. Typical examples of such practices include crop rotation, crop diversification, intercropping, use of improved seed varieties, application of organic fertilizers, and construction of terraces, among others (Tadesse & Ahmed, 2023; Mizik, 2021).

Now then, there is a rich literature on the factors affecting the adoption of CSA practices. Initially, a large set of studies focused on 'technical' determinants: soil fertility, access to a stable source of water, infrastructure (e.g., irrigation, storage facilities, etc.), human capital (e.g., experience, skills, education), extension services and training, dissemination of technological packages, and so on (Tadesse & Ahmed, 2023; Mizik, 2021). Initially, these studies assumed that these technical factors would be sufficient to stimulate CSA; yet they underestimated the complexity of the institutional context within which farmers and agricultural 'systems' operate (Totin et al., 2018). This prompted a new wave of studies looking at an array of socio-economic issues, including access to land (i.e., farm size), other factor/input markets (labor, fertilizers, seed, etc.), output markets, tenure security, associativity (e.g., cooperative membership), and even gender relations (ibid).

Nevertheless, there is one crucial aspect of the agricultural 'system' which has been totally neglected by the literature: *land inequality*. Different from the farm size issue, which relates to the sheer amount of land available to any individual farmer, land inequality refers to the allocation of the total available land across a collection of farmers within a specific geographical area, and in that very concrete sense, it is a *social* phenomenon. Indeed, in developing countries, this allocation is oftentimes historical, political, and institutional, separating it from the production process itself, and reinforcing its 'system' nature (García, 1967; Chonchol, 1996).

Given the above, this study aims to explore how land inequality affects both the likelihood and intensity of adopting CSA practices. For this purpose, a sample of 4,217 farmers from the Peruvian highland region is compiled using the 2022 national agricultural survey (Encuesta Nacional Agropecuaria) of Peru (INEI, 2022). The analysis examines the effect of two alternative measures of land inequality—the Gini and Kakwani indexes—on both the adoption and intensity of use of CSA practices, using a double-hurdle modelling approach.

The results show that land inequality exhibits a marginally significant negative effect on adoption when using the Gini index, and a highly significant negative effect when using the Kakwani index. However, land inequality does not affect the intensity of use at all. Additionally, other factors that positively influence adoption or intensity of use are identified, such as water access, tenure security, extension services, farm income, and credit.

The remaining of this paper is organized as follows: the next section, Section 3.2, establishes the theoretical link between land inequality and CSA adoption and intensity based on a review of the relevant literature; Section 3.3 explains the methodology employed; Section 3.4 presents the data employed in this study; Section 3.5 discusses the results; and Section 3.6 concludes.

3.2. Literature Review

While there is an ample empirical literature on the technical determinants of CSA adoption (and intensity of use) such as soil fertility, access to a stable source of water, infrastructure (e.g., irrigation, storage facilities, transportation, etc.), human capital (e.g., experience, skills, education), extension services and training, dissemination of technological packages, and so on (Tadesse & Ahmed, 2023; Mizik, 2021), only about half of the works look into the market, political, and institutional determinants (Totin et al., 2018). Among them, the literature has focused on access to land (i.e., farm size), other factor/input markets (labor, fertilizers, seed, etc.), output markets, tenure security, associativity (e.g., cooperative membership), and even gender relations (ibid).

However, there are no empirical studies that have examined the relationship between land inequality and the adoption or intensity of CSA practices. Nonetheless, the agricultural economics literature offers clear theoretical insights into their connection, particularly within the context of developing-country agricultural sectors, where land market imperfections play a major role. Specifically, it is argued that land market imperfections lead small farms to face a lower shadow price for labor (e.g., family labor) and higher opportunity costs for land and capital due to credit-rationing, covariate risk, or speculative land values (Besley, 1998; Carter & Mesbah, 1990). On the contrary, large farms face a higher shadow price for labor (due to search and monitoring costs of hired workers), and lower opportunity costs for land and capital. This idea of different relative prices faced by small and large farms (e.g., a minimum and maximum interest rates in capital markets) is elegantly formalized by Allen & Lueck (1998).

As a result of the above, small farms tend to be more intensive in labor than large farms, while large farms tend towards land extensive activities (and capital-intensive techniques—i.e., mechanization) (Ellis, 1993). Because many of the CSA practices are labor-intensive (complex intercropping techniques, composting and manure application, etc.), it is therefore more likely that small farms will adopt them and use them more intensively. This leads Rampa et al. (2020) to argue that a redistribution of land from large to small farms—the reduction of land inequality—increases the probability of adoption and intensity of use of CSA practices, going as far as to propose the concept of a 'climate smart land reform', with land redistribution as one of its main pillars.

Interestingly, the above theory has a strong applicability in the context of this study. Indeed, as revealed by the study data from the 2022 national agricultural survey of Peru (see the data section) Peruvian Andean agriculture is characterized by a very high intensity of labor use relative to land and capital; and many of the CSA practices identified in the survey are, too, labor-intensive—e.g., intercropping, application of organic matter, terracing, etc. (a complete list of CSA practices will be presented later in the data section). Therefore, the Peruvian context presents a suitable environment for exploring the hypothesis that a reduction of inequality enhances the adoption and intensity of use of CSA practices.

3.3. Methodology

In this paper we make use of a 'double-hurdle' model (Cragg, 1971) to explore the impact of land inequality on CSA adoption and intensity of use. This model assumes a two-step data generation process: in the first step or 'first hurdle', farmers decide whether to adopt or not any positive number of CSA practices—captured by a binary probability model. If this hurdle is 'crossed', they decide on the number of CSA practices to adopt (the 'second hurdle')—captured by a truncated count distribution model (Skevas et al., 2022). Furthermore, a central aspect of this model is that the two steps of the data generation process are not constrained to be the same (Cameron & Trivedi, 1998). Hence, the double-hurdle model contains the following two sets of equations and censoring rules:

$$first\ hurdle \begin{cases} d_i^* = \boldsymbol{z_i'}\boldsymbol{\alpha} + \varepsilon_{1,i} \\ d_i = 1\ if\ d_i^* > 0, or\ 0\ if\ d_i^* \leq 0 \end{cases} \quad (1)$$

$$second\ hurdle \begin{cases} y_i^* = \boldsymbol{x_i'}\boldsymbol{\beta} + \varepsilon_{2,i} \\ y_i = y_i^*\ if\ y_i^* > 0\ and\ d_i^* > 0, or\ 0\ otherwise \end{cases} \quad (2)$$

where d_i^* is a latent variable defining the adoption process and d_i is the observed outcome of the binary choice; y_i^* is a latent variable defining the intensity of adaptation or number of CSA practices to adopt and y_i is the observed number of adopted practices; $\boldsymbol{z_i}$ and $\boldsymbol{x_i}$ are vectors of explanatory variables for adoption and intensity, respectively; $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the vectors of parameters to be estimated; and finally, $\varepsilon_{1,i}$ and $\varepsilon_{2,i}$ are the identically and independently normally distributed random errors with zero means and constant variances (Engel & Moffatt, 2014). Given that the $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ parameters do not have a direct interpretation, it is common to estimate the marginal effects at the mean of each variable (Skevas et al., 2022).

McDowell (2003) shows that the log-likelihood function of the double-hurdle model is separable with respect to the parameters, and so it can always be written as the sum of the likelihoods of two different models. Therefore, the log-likelihood function of the double-hurdle model can be maximized by maximizing the two components separately. It is this feature, precisely, which allows us to estimate a binomial probability model (first hurdle) and a truncated count distribution model (second hurdle) separately. Specifically, a 'Logit' model is estimated for the first hurdle. Two alternative models, a zero-truncated Poisson and a zero-truncated negative binomial model, were estimated for the second hurdle, with a test procedure employed to select the most appropriate model[25]. In essence, considering these two options allows us to determine which probability distribution better represents the data generation process.

In relation to the double-hurdle model, two tests must be conducted: first, whether the data-generation process is indeed a two-step ('double-hurdle') process—instead of a one-step process; and whether the two error terms ($\varepsilon_{1,i}$ and $\varepsilon_{2,i}$) are uncorrelated. For the first test, we follow the approach of Tambo & Abdoulaye (2012) and others (e.g., Roco et al. 2014; Skevas et al., 2022) of using a 'likelihood ratio test' (LR) under the null hypothesis that a one-step count model (e.g., a Poisson model) is preferable to a double-hurdle model. Based on the separability and additivity of

---

[25] The test procedure entailed estimating the negative binomial model and checking the statistical significance of the overdispersion parameter, α. If α is found to be statistically significant, it indicates that the negative binomial distribution is more suitable for explaining CSA adoption.

the double-hurdle log-likelihood function discussed in the previous paragraph, the LR test entails estimating the three models under consideration (the Logit and zero-truncated Poisson/negative binomial models of the double-hurdle approach, and the one-step Poisson/negative binomial model), and computing:

$$\lambda = -2(L_{1S} - L_L - L_T) \quad (3)$$

where L_{1S}, L_L, and L_T are the log-likelihood function values of the one-step Poisson/negative binomial, the Logit, and the zero-truncated Poisson/negative binomial models, respectively. On its side, the LR statistic (λ) has a Chi-square distribution with k degrees of freedom (where k is the number of independent variables including a constant). If it exceeds the corresponding critical values at the 1 or 5% levels, the null hypothesis is rejected in favor of the double-hurdle model.

To test for the conditional independence of the error terms, an approach described by Burke et al. (2015) and Skevas et al. (2022) similar to the Heckman test for selection bias is used: first, the Logit model (first hurdle) is estimated and the inverse Mills ratio (IMR) around the probability of adopting CSA practices is predicted; second, the zero-truncated Poisson/negative binomial model (second hurdle) is estimated, including the predicted IMR as an additional control variable. If its coefficient is statistically significant, then the errors of the first and second hurdles are correlated, in which case the empirical design must be reconsidered. If, on the contrary, the IMR coefficient is statistically insignificant, then the errors of the two hurdles are uncorrelated, and the zero-truncated Poisson/negative binomial model is re-estimated without this additional control.

3.4. Data

For this study, the 2022 national agricultural survey ('Encuesta Nacional Agropecuaria', ENA) of Peru (INEI, 2022) is employed. This survey is a nationally representative sample of 30,293 agricultural producers and their productive units ('unidades agropecuarias') across the three main geographical regions of the country (coast, highland, and jungle), as well as across different jurisdictional levels (departments, provinces, and districts). It includes detailed information on production practices and farm finances (e.g., inputs, outputs, and CSA practices[26]), as well as a wide range of socio-economic characteristics of the producers. Due to the specific focus of this study on CSA practices related to crop farming, all farms that did not grow any crops and were dedicated exclusively to livestock activities, as well as all the 'inactive' and 'landless' productive units, were excluded from the sample. Additionally, focus was directed exclusively towards the highland region due to the higher applicability of certain CSA practices (e.g. terracing) in mountainous landscapes. Moreover, the region offers a sufficiently large number of farms by district[27], enabling meaningful calculation of the Gini index. Moreover, to ensure an adequate number of observations for calculating the district-level Gini index, districts with fewer than 30

---

[26] Specifically, the survey contains a chapter that reports on 'buenas prácticas agropecuarias' (good agricultural practices). Although not all of them can be considered CSA practices, several are indeed classified as such by the literature. Other CSA practices such as 'intercropping' are contained in different chapters.

[27] This level of aggregation was chosen because other available levels of aggregation, namely departments and provinces, span across multiple geographical regions. This creates a mismatch between the jurisdictions and what is defined as the highland region.

farms were excluded from the sample[28]. After applying these criteria, the sample process resulted in 4,217 observations.

3.4.1. CSA practices

Nine different CSA practices were identified in the survey. These include: 1) intercropping (‘cultivo asociado’, ‘dispero’, o ‘vergel’); 2) crop rotation; 3) application of organic matter (stubble, manure, compost, humus, etc.); 4) integrated pest control—i.e., the combination of two or more of the following: cultural control, physical control, biological control, ethological control, chemical control (i.e., pesticides)[29], and others; 5) improved seed (‘semilla certificada’); 6) soil analysis and recovery (of eroded/compacted/saline soils); 7) water management (water analysis, measurement, timing, etc.); 8) terracing; and 9) contouring.

As mentioned earlier, most of these CSA practices are labor-intensive. For example, complex intercropping combinations such as the traditional ‘milpa’ and ‘conuco’ combinations

---

[28] Although this choice may seem somewhat arbitrary, a minimum sample size of 30 is generally used as a rule of thumb in statistical analysis based on the Central Limit Theorem (Hogg et al., 2014).

[29] The use of pesticides as a CSA measure is highly debated. On the one hand, it may be considered as a CSA practice ever since it contributes to a higher productivity and adaptation (two of the three pillars of CSA). However, an intensive use of chemical pesticides may be in direct conflict with the mitigation objective—the third pillar (Mizik, 2021; Tambet & Stopnitzky, 2021). The same situation applies to other ‘green revolution’ technologies, such as fertilizers. Given this lack of consensus, we have focused exclusively on practices that are unambiguously recognized by the literature as CSA practices. In the particular case of integrated pest control, the use of chemical pesticides is allowed as a CSA practice when used in combination with other cultural, physical, biological, or ethological control practices, among others.

require careful handling by the farmer[30], and the manual recollection of manure from the fields and its application, or the construction of terraces on the steep slopes of the Andes, require intense and long working hours (Chonchol, 1996). Therefore, the Peruvian highland context is ideal for testing the main hypothesis, which posits that a higher labor-intensity of CSA practices creates an inverse relationship between land inequality and CSA adoption and intensity of use.

3.4.2. Land inequality

Aligned with the main research objective, two different measures of land inequality are included as independent variables, the Gini and Kakwani indexes. On one side, the construction of the *Gini* coefficient involves arranging farmers according to their land endowments and constructing a Lorenz curve, which illustrates the cumulative distribution of land compared to a line of perfect equality. Then, the index is calculated as the ratio of the area between the Lorenz curve and the line of perfect equality to the total area under the latter, yielding a value between 0 and 1—with higher values indicating greater inequality (Luebker, 2010). This coefficient is calculated from the farm-level data at the district level, and then applied to each farm individually.

On the other side, the *Kakwani* index (Wang et al., 2022) enables the measurement of inequality at the household or farm level through the concept of 'relative deprivation'. Specifically, it compares the land endowments of each productive unit to the land endowments of the rest of the

---

[30] The customary *milpa* is the association of three nutritional elements—maize, beans, and calabashes—that together conform a single technological arrangement: the maize provides a support structure for the beans, while the abundant foliage of the calabashes protects the soil from overheating and from the impact of the rain. Other complex combinations like the *conuco* include a wider diversity of crops (Chonchol, 1996).

farms in the sample: the higher the index for a particular unit, the higher its relative deprivation, or the more unfavorable its social position when compared to others. Its calculation proceeds as follows: first, the $N = 1, 2, \dots, n$ farms in the sample are rearranged according to their land endowments, X, in ascending order, i.e., $x_1 \leq x_2 \leq \cdots \leq x_n$. Then, the relative land deprivation for the i'th productive unit is:

$$Kakwani_i = \frac{1}{N\mu_x} \sum\nolimits_{j=i+1}^{n} (x_j - x_i)$$

Or, equivalently,

$$Kakwani_i = \gamma_{x_i}^{+} \left[\frac{(\mu_{x_i}^{+} - x_i)}{\mu_x} \right] \quad (4)$$

where μ_x is the average cultivated land endowment of the sample, $\mu_{x_i}^{+}$ is the average cultivated land endowment of those farms with more land than x_i, and $\gamma_{x_i}^{+}$ is the proportion of farms with more land than farm i (i.e., M/N, where $M = j, \dots, n$). This index is bounded between zero and one. The higher the proportion of farms with more land than farm *i*, or the larger their average land compared to farm *i*'s endowment, the closer to 1 the index will be, and the smaller the relative land endowments of farm *i* compared to the rest of the productive units in the sample. Conversely, the closer to zero, the lower the relative deprivation of the farm, and the larger its land endowments compared to the rest of productive units.

Aligned with the research hypothesis, negative coefficients are anticipated for both the *Gini* and *Kakwani* indexes in both hurdles. These results will indicate that the lower level of inequality or of relative land deprivation, the higher the adoption and intensity of use of CSA practices.

3.4.3. Other explanatory variables

Besides land inequality, a set of additional explanatory variables stemming from the literature (refer to Mizik [2021] for a comprehensive review of the determinants of CSA adoption) are included in the analysis. These variables include the years of *age* of the main producer; their educational attainment—captured by two dummy variables, *low educ* and *high educ*[31]; the number of years of farming *experience*; *farm income* (measured in Soles); *farm size* (measured in hectares); *tenure security*, captured by the percentage of hectares owned by the producer[32]; and a set of six dummy variables. These dummy variables signify whether the main producer is a *female*; whether the farm has stable *water access* (river, spring, well, dam, reservoir, or others); whether the main producer received training from *extension* services; whether the productive unit had access to *credit*; whether the main producer belonged to any *association* (cooperatives, committees, etc.); and whether the main producer had *off-farm income*.

Finally, because the nine CSA practices considered in this work relate exclusively to crop-growing activities and are not applicable to livestock farming, it becomes necessary to control for

---

[31] The original ENA survey measures education in ten discrete levels. For parsimony of the model, we regroup them into these two categories, where *loweduc* includes no level, initial education, incomplete primary education, complete primary education, and incomplete secondary education; and *higheduc* includes complete secondary education, incomplete superior non-university education, incomplete superior university education, complete superior non-university education, and complete superior university education.

[32] The ENA survey documents five types of tenure arrangements: owned land, rented land, communal land, share-holding, and others. According to the literature (Demsetz, 1967; FAO, 2002), individual property rights (as opposed to other forms of tenure) enhance tenure security, and so we focus exclusively on this category.

the percentage of *crop output* over total output of the farm. The full variable list and descriptive statistics are presented in Table 3.1.

Table 3.1. Descriptive statistics

VARIABLES	unit	mean	s.d.	min.	max.
Gini	index (0–1)	0.567	0.120	0.202	0.948
Kakwani	index (0–1)	0.828	0.145	6.11e-05	0.999
age	years	55.96	14.53	18	98
high educ	dummy (base: low educ)	0.232	0.422	0	1
experience	years	29.61	14.55	1	82
farm income	Soles	7,946	29,143	0	705,600
farm size	hectares	3.378	32.26	0.005	2,000
tenure security	percentage	73.68	40.73	0	100
female	dummy (base: male)	0.346	0.476	0	1
water source	dummy (base: no water source)	0.537	0.499	0	1
extension	dummy (base: no extension)	0.054	0.226	0	1
credit	dummy (base: no credit)	0.077	0.267	0	1
association	dummy (base: no association)	0.036	0.186	0	1
off-farm income	dummy (no off-farm income)	0.430	0.495	0	1
crop output	percentage	66.78	32.22	0.067	100

Most of the above explanatory variables are expected to positively impact CSA adoption and intensity of use. For example, *high educ* is expected to increase farmer's awareness of climate change and make them more receptive towards new practices (Branca & Perelli, 2020); a stable *water source* is a pre-requisite for sustainable agriculture (Mango et al., 2018); *farm income* and *credit* may provide the necessary means to conduct certain CSA practices that require heavy investments (e.g., contouring, terracing, etc.) (Navarro-Castañeda et al., 2021); and *tenure security* may create the necessary incentives for those investments (Demsetz, 1967; Zegeye, 2021). Yet, there are at least three factors which have been found in previous studies to have a negative effect.

In particular, Roco et al. (2014) and Thinda et al. (2020) document that older farmers (*age*) are less open to new practices and thus are less likely to adopt climate change adaptation strategies; also, *off-farm income* has been found to have a negative influence, possibly due to the replacement of agriculture as the main source of income of the household (Abegunde et al., 2020); and finally, a clear constraint on *female*-led productive units has been documented, given the prevailing gender roles and imbalances of power between genders in some traditional contexts (Murray et al., 2016; Tsige et al., 2020; Navarro-Castañeda et al., 2021).

3.5. Results

3.5.1 Use of CSA practices

First, we begin by exploring the percentage of farmers who adopted any positive quantity of CSA practices—the first hurdle—illustrated in Figure 1; the percentage of farmers adopting different positive quantities of CSA practices or intensity of use—the second hurdle—illustrated in Figure 2; and the percentage of farmers using each of the identified CSA practices, in Figure 3.

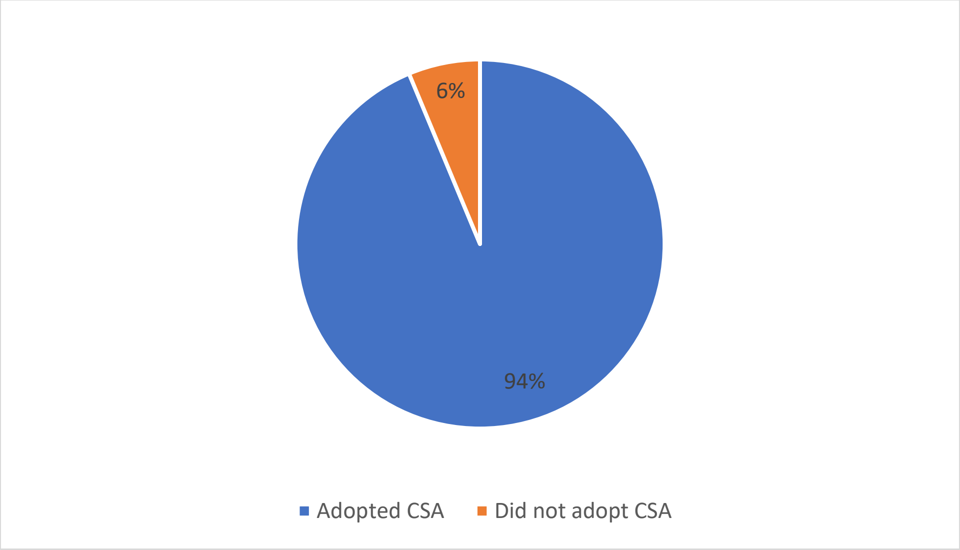


Figure 3.1. Percentage of farmers adopting CSA practices

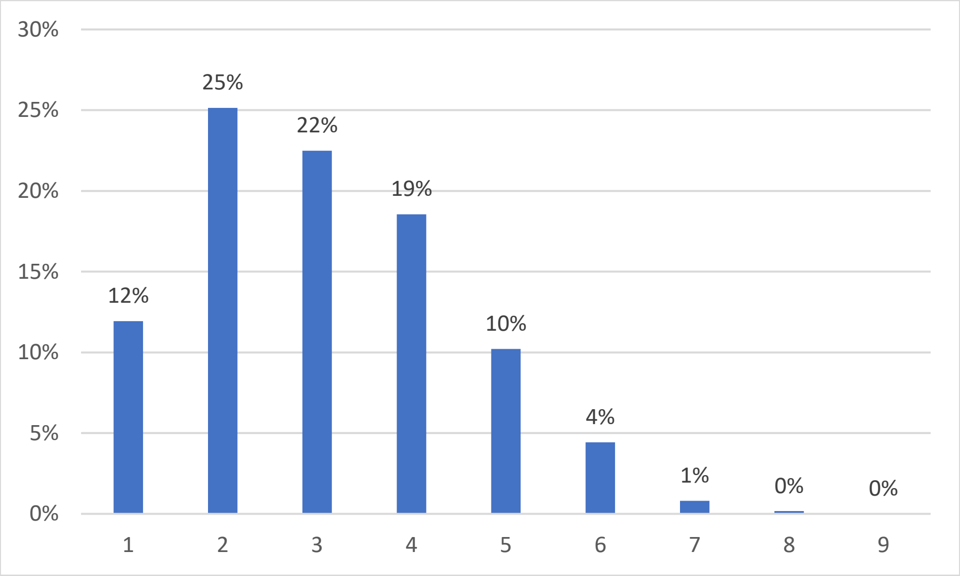


Figure 3.2. Intensity of use of CSA practices

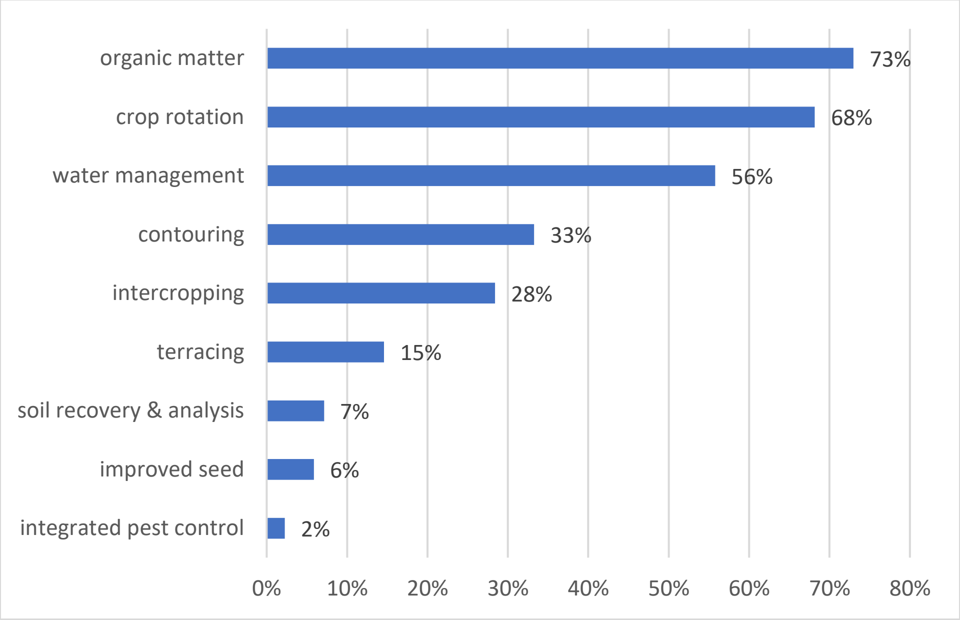


Figure 3.3. Percentage of farmers using different CSA practices.

In Figure 1, we see that the majority of farmers in the sample (94%) adopted at least one CSA practice, meaning that they ‘crossed’ the first hurdle. With regards to the second hurdle, evidenced in Figure 2, we see that the distribution of the intensity of use is highly skewed to the right; or, in other words, that the majority of farmers adopted only a few CSA practices. Specifically, the mean number of adopted practices among the farmers who crossed the first hurdle was 3.1, and less than 1% of the farmers adopted seven or more practices. Finally, among the CSA practices, the most common practice was the application of organic matter (73% of the farmers), followed by crop rotation (68%), and water management (56%).

These results are somewhat similar to those found in different contexts. For instance, looking at the adoption of nine different adaptation practices among land reform beneficiaries in South Africa, Thinda et al. (2020) find a mean number of adopted practices of 3.6, although the distribution of the number of adopted practices (the intensity of use) is not that skewed. Yet, adoption of adaptation measures wasn’t as high in other contexts of study, such as that of central Chile (Roco et al., 2014), where the adoption rate was of about 57%, with an average intensity of use of only 1.74 (among a total of 11 practices). Similarly, Skevas et al. (2022) find among a sample of Missouri farmers an adoption rate of only 66%, with a mean of 2.2 adaptation strategies among the adopting farmers. Therefore, the rates of adoption and intensity of use seem to be highly context-specific, as this diversity of finding suggests.

3.5.2 Decision and intensity of CSA adoption

Before diving into the estimation results, let's first explore the steps taken to confirm the appropriateness of the selected modeling approach. The first step in the estimation procedure involved determining which distribution was more appropriate for the data generation process between the Poisson and the negative binomial distributions. It was found that the overdispersion parameter, α, was not statistically different from zero (the likelihood-ratio Chi-square test that α equals zero is 0.00, with a p-value of 1.000), regardless of the choice of the inequality index (the Gini or Kakwani indexes) used as main independent variable (the zero-truncated negative binomial regression outputs are presented in Table A4 in the Appendix). Hence, the Poisson distribution is preferred over the negative binomial distribution to model the outcome of CSA practices used by the sample farmers.

The second step involved checking whether a double-hurdle specification was preferred over a one-step count model using the LR test of equation (3). It was found that the LR test statistic lambda (λ) exceeded the appropriate Chi-square critical value (at the 1% level) with 17 degrees of freedom in all the estimations (see Appendix, Table A5), thus strongly rejecting the null hypothesis that the one-step count model outperforms the double-hurdle model. The third, and final step, was to check whether the error terms of the two hurdles were uncorrelated. It was found that the IMR coefficient was statistically insignificant at the conventional levels using either the Gini or Kakwani indexes (p-values of 0.164 and 0.165, respectively), implying that the random errors of the first and second hurdles were uncorrelated. Hence, it was concluded that the empirical design was appropriate for our case study, and so we focused the discussion on the double hurdle model.

As an additional assessment, we checked for any potential multicollinearity issues in the data (see a correlation matrix in Table A3 in the Appendix), finding no problems among the variables. The regression outputs for the first and second hurdle models using the Gini and Kakwani indexes and including the marginal effects evaluated at the mean of each variable are presented in tables 3.2 and 3.3.

Table 3.2. First-hurdle Logistic regression output

VARIABLES	**Gini** *d*	**Gini** mg. effect	**Kakwani** *d*	**Kakwani** mg. effect
Gini	-1.094*	-0.027*		
	(0.598)	(0.015)		
Kakwani			-1.271**	-0.031**
			(-0.535)	(0.013)
age	0.011	0.000	0.011	0.000
	(0.008)	(0.000)	(-0.008)	(0.000)
high educ	0.171	0.004	0.173	0.004
	(0.175)	(0.004)	(-0.176)	(0.004)
experience	-0.004	-0.000	-0.004	-0.000
	(0.007)	(0.000)	(-0.007)	(0.000)
farm income	3.16e-07	7.66E-09	0.000	-7.71E-08
	(4.49e-06)	(1.09E-07)	(0.000)	(8.98E-08)
farm size	0.002	0.000	-0.002	-0.000
	(0.006)	(0.000)	(-0.002)	(0.000)
tenure security	-0.001	-0.000	0.000	-3.48E-06
	(0.002)	(0.000)	(-0.002)	(0.000)
female	0.026	0.001	0.065	0.002
	(0.143)	(0.003)	(-0.144)	(0.003)
water source	3.337***	0.081***	3.363***	0.081***
	(0.310)	(0.007)	(-0.312)	(0.007)
extension				
credit	0.469	0.011	0.469	0.011
	(0.362)	(0.009)	(-0.362)	(0.009)
association	-0.535	-0.013	-0.596	-0.014
	(0.372)	(0.009)	(-0.371)	(0.009)
off-farm income	0.058	0.001	0.070	0.002
	(0.147)	(0.003)	(-0.147)	(0.004)
crop output	-0.007***	-0.000***	-0.008***	-0.000***
	(0.002)	(0.000)	(-0.002)	(0.000)
Pseudo R2	0.172		0.173	
Prob > Chi2	0.000		0.000	
Log-likelihood	-807.107		-805.900	
Observations	3,989		3,989	

Constants omitted. Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.3. Second-hurdle zero-truncated Poisson regression output

	Gini	**Gini**	**Kakwani**	**Kakwani**
VARIABLES	*y*	mg. effect	*y*	mg. effect
Gini	0.066	0.179		
	-0.084	0.227		
Kakwani			0.075	0.203
			-0.080	0.218
age	0.002	0.004	0.002	0.004
	-0.001	0.003	-0.001	0.003
high educ	0.048*	0.131*	0.048*	0.129*
	-0.025	0.068	-0.025	0.068
experience	0.002	0.004	0.002	0.005
	-0.001	0.003	-0.001	0.003
farm income	7.65e-07***	2.08E-06***	7.78e-07***	2.11E-06***
	-2.7E-07	7.4E-07	-2.7E-07	7.4E-07
farm size	-1.8E-04	-4.9E-04	-9.2E-05	-2.5E-04
	-3.3E-04	9.1E-04	-3.3E-04	8.8E-04
tenure security	0.001***	0.002***	0.001***	0.002***
	0.000	0.001	0.000	0.001
female	0.018	0.049	0.014	0.039
	-0.022	0.059	-0.022	0.060
water source	0.649***	1.762***	0.648***	1.759***
	-0.024	0.059	-0.024	0.059
extension	0.151***	0.410***	0.151***	0.410***
	-0.039	0.106	-0.039	0.106
credit	0.092***	0.251***	0.092***	0.251***
	-0.035	0.095	-0.035	0.095
association	0.010	0.026	0.013	0.036
	-0.053	0.143	-0.053	0.144
off-farm income	0.032	0.087	0.030	0.082
	-0.023	0.061	-0.023	0.062
crop output	0.000	-0.001	0.000	-0.001
	0.000	0.001	0.000	0.001
Pseudo R2	0.0747		0.0748	
Prob > chi2	0.0000		0.0000	
Log-likelihood	-6347.459		-6347.335	
Observations	3,952		3,952	

Constants omitted. Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Starting with adoption decisions (first hurdle model of Table 3.2), we find that both the Gini coefficient and the Kakwani land deprivation index have a negative and significant marginal effect (at the 10% and 5% levels, respectively), confirming our hypothesis that land inequality translates into a lower adoption probability of CSA practices. Concretely, on average, a one-percentage point increase in the Gini coefficient decreases the probability of adoption approximately by 2.7 percent, while a one percentage-point increase in the Kakwani land deprivation index decreases the probability of adoption by roughly 3.1 percent.

Among the set of control variables, only two have very significant marginal effects, regardless of the use of Gini or Kakwani: on the one hand, *water source* has a robust positive effect (at the 1% level), indicating that farms with a stable source of water (e.g., river, spring, well, etc.) are about 8.1 percent more likely to adopt CSA practices than farms located in exclusively rain-fed areas. Indeed, having a stable source of water is a pre-requisite for sustainable agriculture (Mango et al., 2018; Kifle, 2021; Sertse et al., 2021); and many of the CSA practices contemplated in this study (e.g., contouring, terracing, soil recovery and analysis, improved seed, etc.) require considerable investments (Navarro-Castañeda et al., 2021), which may be riskier to undertake in environments with unpredictable weather patterns.

On the other hand, *crop output* has a very significant negative effect on the adoption probability. This result seems counter-intuitive, though, since all of the CSA practices examined here relate exclusively to cropping activities (as opposed to other agricultural activities like livestock raising), and therefore one would expect that as crop output is more preponderant, the higher the probability of adoption. Nonetheless, as noted by Kifle et al. (2022), a negative coefficient could be indicative of the importance of diversification strategies in the generation of alternative sources of income and the reduction of risk, which potentially, could result in a higher probability of adoption. Finally, we note that the *extension* variable has been omitted from the first hurdle regression since it predicts adoption perfectly.

Passing to the intensity decisions (second hurdle model of Table 3.3), our results show that neither the Gini nor the Kakwani indexes have any effect on the intensity of adoption of CSA practices, as the agricultural economics theory presented in the literature review section would suggest. However, there is a larger set of control variables that have a positive influence on the intensity decision: to begin with, *high educ* has a marginally significant effect (i.e., at the 10% level) on the number of practices adopted. Concretely, producers with high levels of education (i.e., from complete secondary education onwards) adopt around 13 percent more practices than producers with low levels of education, which is in line with the previous literature on CSA adoption and intensity of use, according to which more educated farmers are more aware of the importance of the three objectives of CSA—increasing productivity, adapting to climate change, and mitigating greenhouse gas emissions—and are also more receptive to new farm practices that contribute towards those goals (Branca & Perelli, 2020; Gebru et al., 2020; Neway & Zegeye, 2022).

Additionally, we see that *farm income* has a very significant (i.e., at the 1% level) positive effect on intensity of use, confirming the well-established notion that income plays a major role in the adoption of CSA practices through a higher ability to invest and cope with risk (Abegunde et al., 2020; Mango et al., 2018). In the same sense, *credit* also has a positive and very significant effect; once more, hinting at the importance of the capacity to invest in such practices (Zerssa et al., 2021; Van Dijk et al., 2020). Similarly, *tenure security* also has a very significant positive effect on the intensity of use, which is in line with the literature on property rights stressing that increased tenure security permits the internalization of costs and benefits and, consequently, incentivize investments (Coase, 1960; Demsetz, 1967). This is also confirmed by studies on CSA adoption and intensity of use, such as those by Teklewold et al. (2019), Zegeye (2021), and Bedeke et al. (2019).

Finally, we see that *water source* has a very robust positive effect on the intensity of use due to the vitality of this resource for agricultural production (Mango et al., 2018; Kifle, 2021;

Sertse et al., 2021); and *extension*, this time, is not dropped from the regression, exhibiting a large and very significant effect on the intensity of use. This denotes the crucial role that extension services play in the promotion of CSA practices, as has been found in Ghana (Akrofi-Atitianti et al., 2018), India (Khatri-Chhetri et al., 2019), Mali (Andrieu et al., 2017), and several other places. Contrary to other studies, however, we do not find any effect of being a *female* (e.g., Murray et al., 2016; Fentie & Beyene, 2019), of belonging to any form of *association* (e.g., Makate, 2019; Bedeke et al., 2019), or of having *off-farm income* (e.g., Setshedi & Modirwa, 2020; Maru et al., 2022).

3.6. Conclusions

This study used a double hurdle modelling approach to analyze the impact of land inequality on the likelihood of adopting and the intensity of use of climate smart agriculture (CSA) practices. The empirical application used a sample of 4,217 farmers from the Peruvian highland region and explored two different inequality measures: the Gini and the Kakwani indexes.

The results show that the vast majority of sampled farmers (94%) have adopted at least one CSA practice. Among those who have adopted CSA practices, the average intensity of use, or the number of practices adopted, was three. Among the nine practices examined in this study, the most commonly adopted CSA practices were the application of organic matter, followed by crop rotation, and water management. However, other important CSA practices such as soil recovery and analysis, improved seed, and integrated pest control, were adopted by a relatively small portion of the sampled farmers. These findings highlight the opportunity for extension specialists and policy makers to concentrate their efforts on promoting these practices further, thereby enhancing overall CSA adoption rates.

The results further highlight the adverse effect of land inequality on the adoption of CSA practices. Specifically, both the Gini coefficient and the Kakwani relative land deprivation index demonstrate a negative effect on CSA adoption, although they did not affect the intensity of use of CSA practices. Consequently, reducing land inequality could potentially promote the adoption of CSA practices in the Peruvian highlands, lending support to the concept of 'climate smart rural reform' proposed by Rampa et al. (2000), where mitigating land inequality holds central significance.

Apart from land inequality, several other factors are shown to significantly impact both the adoption and/or intensity of CSA practices. Specifically, access to a reliable water source, tenure security, extension services, farm income, and credit access, emerge as crucial determinants that facilitate the adoption and/or intensity of use of CSA practices. Therefore, efforts to promote climate smart agriculture in the region should encompass initiatives to strengthen these factors, alongside measures to reduce land inequality.

It is important to highlight a few considerations that this study has not addressed, which could serve as valuable avenues for future research. Firstly, institutional changes typically unfold over extended periods, impacting economic outcomes gradually (Acemoglu et al., 2001). Consequently, the effects of a reduction of land inequality—viewed here as an 'institutional' change—may manifest with a delay in influencing CSA adoption and intensity of use. Future research could use a longer time frame to analyze the long-run effects of land inequality on CSA adoption. Secondly, this research on the relationship between land inequality and CSA adoption and intensity of use focuses on a single region of Peru. Future studies could expand upon this work in different ways, such as by investigating whether similar findings would be observed in other regions of Peru or in other developing countries, where cultural and institutional differences may impact levels of land inequality differently. Additionally exploring different time periods or types of farms could offer further insights into this relationship.

GENERAL CONCLUSION

In this Dissertation, we have explored the effect of economic inequality—particularly *land* inequality—on three different aspects of agricultural production in the context of developing country agriculture: technical inefficiency, total factor productivity, and the probability of adoption and intensity of use of climate-smart agricultural practices.

We have departed from a common theoretical proposition: that according to which factor market imperfections in developing-country agriculture cause small and large farms to face differing shadow prices for land, capital, and labor, causing massive macroeconomic technical inefficiency, lower average total factor productivity, and a higher adoption probability of climate-smart agricultural practices by small farms—justifying the idea of a 'climate-smart rural reform'.

Our three papers demonstrate each of these points, using the best information and methods available to us. Specifically, in the first paper, we have used a panel of fifteen Latin American rural sectors over a period of seventeen years, and implemented a Stochastic Frontier Analysis approach, finding a very significant *positive* effect of rural income inequality on technical *in*efficiency. In the second paper, we have used a sample of almost three-hundred transitory-crop farmers from the Peruvian highlands, and tested the effect of land inequality (as measured by the Gini, Theil, and Kakwani indexes) on total factor productivity using system GMM and a control function approach, finding a significant negative influence of both the Gini and the Theil coefficients. And in the third paper, we have used a wider sample of over four-thousand farmers from the same region to analyze the effect of land inequality on the probability of adoption and intensity of use of CSA practices through the use of a double-hurdle model, finding a significant negative impact on adoption.

Altogether, the findings from these three papers conform to the theoretical framework that inspires them, and stress the need to mitigate economic inequality in developing-country agricultural sectors. Typically, this initiative must be accompanied by parallel efforts to strengthen other technical and socioeconomic factors, such as water access, tenure security, extension services, access to credit, and farm income. Thus, the strategy for rural development cannot be centered in one single aspect—e.g., land inequality—but must have a holistic approach. Yet, we have shown, land inequality plays a major strategic role, although the avenues of *how* that can be accomplished have not been explored in this Dissertation, and remain a crucial task for consecutive research efforts.

As a general guideline, future research should also focus on unveiling the precise mechanisms behind the studied relationships. Indeed, in this Dissertation we have been able to confirm the predicted signs of those relationships, but not the internal or external forces that create them; and so far, we can only rely on the theoretical explanations provided by the agricultural economics literature. These theoretical propositions need to be addressed more thoroughly, through methodologies that permit, for instance, the identification differing shadow prices for factor-inputs and their consequent effects on input intensity, efficiency, productivity, and the adoption of climate-smart agricultural practices. Effectively, doing so would allow us to better understand the phenomena at hand, and to propose adequate and more specific solutions.

Another general stream of research could focus on the dynamic implications of these findings. Indeed, in each of the three essays of this Dissertation we have focused on short-run economic outcomes (inefficiency, productivity, and probability of adoption) and, at best, compared productivity at one point in time with productivity at a later stage. But even doing so is not, by any means, a ‘dynamic’ analysis. This would entail, by definition, analyzing the impact of inequality on aspects such as capital *accumulation*, economic *growth*, total factor productivity *change*, the *take-up rate* of climate-smart agricultural practices, and so on. This could heighten our understanding of the overall development process of developing countries.

Finally, there is the issue of external validity of our studies. Indeed, our three studies are conducted in very specific contexts—Latin American rural sectors and Peruvian Andean agriculture—but it would be worthwhile to examine if the studied relationships are present in other institutional contexts and agricultural systems. This would give more weight to the proposed theoretical framework, allowing it to serve as a basis of a development strategy for the wider community of developing countries.

We conclude by saying that the evidence presented in this Dissertation is by no means exhaustive or conclusive; rather, it is meant to re-awake interest in the pressing matter of economic inequality—and especially, land inequality—in developing countries. This subject was the center of heated academic discussions during the 50s, 60s, and 70s, but with the introduction of the so-called neoliberal transformations of the following decades, it became largely relegated and excluded from the academic and political debates. Our goal has been to revive those discussions by giving them a new fresh air, reassessing them with the use of novel methodologies and richer data sets that were not available during those times, and merging them with the pressing issues of the present, such as climate change. We hope, at least, to have contributed to that objective; and at most, to have pushed the frontier of knowledge even if by ‘millimetric’ proportions.

REFERENCES

1. Abegunde, V. O., Sibanda, M., & Obi, A. (2020). Mainstreaming climate-smart agriculture in small-scale farming systems: a holistic nonparametric applicability assessment in South Africa. *Agriculture*, *10*(3), 52.
2. Acemoglu, D. (1996). A microfoundation for social increasing returns in human capital accumulation. *The Quarterly Journal of Economics*, 111(3), pp. 779–804. https://doi.org/10.2307/2946672
3. Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5), 1369-1401.
4. Ackerberg, D., Caves, K., & Frazer, G. (2006). *Structural identification of production functions*. (MPRA Paper No. 38349). https://mpra.ub.uni-muenchen.de/38349/
5. Adamopoulos, T. (2008). Land inequality and the transition to modern growth. *Review of Economic Dynamics*, 11(2), pp. 257–282. https://doi.org/10.1016/j.red.2007.07.004
6. Ahluwalia, M. S. (1976). Inequality, poverty and development. Journal of Development Economics, 3(4), pp. 307–342. DOI: 10.1016/0304-3878(76)90027-4
7. Aigner, D. J, Lovell, C. K. & Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6 (1), 21–37. https://doi.org/10.1016/0304-4076(77)90052-5
8. Akrofi-Atitianti, F., Ifejika Speranza, C., Bockel, L., & Asare, R. (2018). Assessing climate smart agriculture and its determinants of practice in Ghana: A case of the cocoa production system. *Land*, *7*(1), 30.
9. Alauddin, M., Headey, D., & Rao, D. S. P. (2005). *Explaining agricultural productivity levels and growth: An international perspective*. (CEPA, Working Paper Series No. 02/2005). https://economics.uq.edu.au/files/5322/WP022005.pdf

10. Alesina, A., & Rodrik, D. (1994). Distributive politics and economic growth. The Quarterly Journal of Economics, 109(2), pp. 465–490. https://doi.org/10.2307/2118470
11. Allen, D. W., & Lueck, D. (1998). The nature of the farm. *The Journal of Law and Economics*, 41(2), pp. 343–386. https://doi.org/10.1086/467393
12. Anderson, T. W., & Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18(1), pp. 47–82. https://doi.org/10.1016/0304-4076(82)90095-1
13. Andrieu, N., Sogoba, B., Zougmore, R., Howland, F., Samake, O., Bonilla-Findji, O., ... & Corner-Dolloff, C. (2017). Prioritizing investments for climate-smart agriculture: Lessons learned from Mali. *Agricultural Systems*, *154*, 13-24.
14. Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), pp. 277–297. https://doi.org/10.2307/2297968
15. Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), pp. 29–51. https://doi.org/10.1016/0304-4076(94)01642-D
16. Asefa, S., 2011. *Analysis of technical efficiency of crop producing smallholder farmers in Tigray, Ethiopia*. (MPRA Paper No. 40461). https://mpra.ub.uni-muenchen.de/40461/
17. Banerjee, A., & Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *American Economic Review*, 95(4), pp. 1190–1213. DOI: 10.1257/0002828054825574
18. Battese, G. E. & Coelli, T. J., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20 (2), 325–332. https://doi.org/10.1007/BF01205442

19. Bauluz, L., Govind, Y., & Novokmet, F. (2020). *Global land inequality*. (WID.world Working Paper N° 2020/10). https://shs.hal.science/halshs-03022360/
20. Bedeke, S., Vanhove, W., Gezahegn, M., Natarajan, K., & Van Damme, P. (2019). Adoption of climate change adaptation strategies by maize-dependent smallholders in Ethiopia. NJAS-Wageningen Journal of Life Sciences, 88, 96-104.
21. Bejarano, J. A. (1997). *Colombia: inseguridad, violencia y desempeño económico en las áreas rurales*. Bogotá: Universidad Externado de Colombia.
22. Belotti, F., Daidone, S., Ilardi, G. & Atella, V., 2013. Stochastic frontier analysis using Stata. *The Stata Journal*, 13 (4), 719–758. https://doi.org/10.1177/1536867X1301300404
23. Benabou, R. (1996). Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance. *The American Economic Review*, pp. 584-609. https://www.jstor.org/stable/2118213
24. Bendig, D., & Hoke, J. (2022). Correcting Selection Bias in Innovation and Entrepreneurship Research: A Practical Guide to Applying the Heckman Two-Stage Estimation. *Available at SSRN 4105207*.
25. Benhabib, J., & Rustichini, A. (1996). Social conflict and growth. *Journal of Economic Growth*, 1, pp. 125–142. https://doi.org/10.1007/BF00163345
26. Berry, R. A., & Cline, W. R., 1979. *Agrarian structure and productivity in developing countries*. Baltimore: Johns Hopkins University Press.
27. Bertola, G. (1993). Market structure and income distribution in endogenous growth models. *American Economic Review*, 83(2), pp. 1184–99.
28. Besley, T. J., 1998. How do market failures justify interventions in rural credit markets? *In* C. K., Eicher and J. M. Staatz, eds. *International Agricultural Development*, 3rd edition. Baltimore: Johns Hopkins University Press, 370–389.

29. Berry, R. A., & Cline, W. R. (1979). *Agrarian structure and productivity in developing countries*. Baltimore: Johns Hopkins University Press.
30. Bhalla, S. S., & Roy, P. (1988). Mis-specification in farm productivity analysis: the role of land quality. *Oxford Economic Papers*, 40(1), pp. 55–73. https://doi.org/10.1093/oxfordjournals.oep.a041846
31. Binswanger, H. P., 1987. Impact of government policies and legal provisions on land use, land value and agrarian structure: analytical framework, hypotheses and terms of reference. *World Bank Mimeo*, Washington, DC.
32. Birdsall, N., & Londoño, J. L. (1997). Asset inequality matters: an assessment of the World Bank's approach to poverty reduction. *The American Economic Review*, 87(2), pp. 32–37. https://www.jstor.org/stable/2950879
33. Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), pp. 115–143. https://doi.org/10.1016/S0304-4076(98)00009-8
34. Borrero, H., & Garza, N. (2019). Growth and distribution endogenously determined: a theoretical model and empirical evidence. *Brazilian Journal of Political Economy*, *39*, 344-361.
35. Brada, J. C., & Bah, E. M. (2014). *Growing income inequality as a challenge to 21st century capitalism*. (Italian Association for the Study of Economic Asymmetries, Working Paper No. 1402) https://EconPapers.repec.org/RePEc:ais:wpaper:1402
36. Bogdanski, A. (2012). Integrated food–energy systems for climate-smart agriculture. Agriculture & Food Security, 1, 1-10.
37. Branca, G., & Perelli, C. (2020). ‘Clearing the air’: common drivers of climate-smart smallholder food production in Eastern and Southern Africa. *Journal of Cleaner Production*, *270*, 121900.

38. Burke, W. J., Myers, R. J., & Jayne, T. S. (2015). A triple-hurdle model of production and market participation in Kenya's dairy market. *American Journal of Agricultural Economics*, *97*(4), 1227-1246.
39. Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53). Cambridge university press.
40. Carter, M. R., 1984. Identification of the inverse relationship between farm size and productivity: an empirical analysis of peasant agricultural production. *Oxford Economic Papers*, (36), 131–45. https://doi.org/10.1093/oxfordjournals.oep.a041621
41. Carter, M. R. & Kalfayan, J., 1989. *A general equilibrium exploration of the agrarian question*. (Department of Agricultural Economics, University of Wisconsin-Madison, Staff Paper No. 279).
42. Carter, M. R., & Mesbah, D., 1990. *Economic theory of land markets and its implications for the land access of the rural poor.* (LTC Paper No. 141).
43. Cendales, A., N. Garza & Arcila, A., 2022, Political machines and the curse of public resources in subnational democracies. *Journal of Economic Studies*, 49 (4), 665–682. DOI: 10.1108/JES-03-2021-0148
44. Chenery, H., Ahluwalia, M. S., Bell, C. L. G., Duloy, J. H., & Jolly, R. (1976). *Redistribución con crecimiento*. Madrid: Editorial Tecnos, S.A.
45. Chonchol, J. (1996). *Sistemas agrarios en América Latina. De la etapa prehispánica a la modernización conservadora*. Fondo de Cultura Económica, Santiago de Chile.
46. Coase, R. H. (1960). The problem of social cost. *The journal of Law and Economics*, 3, pp. 1–44. https://doi.org/10.1086/466560
47. Coelli, T. J., Rao, D. S. P., O'Donnell, C. J. & Battese, G. E., 2005. *An introduction to efficiency and productivity analysis*. New York: Springer.

48. Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: journal of the Econometric Society*, 829-844.
49. Craig, B. J., Pardey, P. G., & Roseboom, J. (1997). International productivity patterns: Accounting for input quality, infrastructure, and research. *American Journal of Agricultural Economics*, 79(4), pp. 1064–1076. https://doi.org/10.2307/1244264
50. Deininger, K. & Feder, G., 2001. Land institutions and land markets. *In* B. L. Gardner and G. C. Rausser, eds. *Handbook of Agricultural Economics*, 1A. Amsterdam: North-Holland, 287–331.
51. Ellis, F., 1993. *Peasant economics: Farm households and agrarian development* (2nd Ed.). Cambridge: Cambridge University Press.
52. Deininger, K., & Binswanger, H. (1999). The evolution of the World Bank's land policy: principles, experience, and future challenges. *The World Bank Research Observer*, 14(2), pp. 247–276. https://doi.org/10.1093/wbro/14.2.247
53. Deininger, K., & Olinto, P. (1999). *Asset distribution, inequality, and growth*. (SSRN). https://ssrn.com/abstract=630745
54. Deininger, K., & Squire, L. (1998). New ways of looking at old issues: inequality and growth. *Journal of Development Economics*, 57(2), pp. 259–287. https://doi.org/10.1016/S0304-3878(98)00099-6
55. Demsetz, H. (1967). Towards a Theory of Property Rights. The American Economic Review, 57(2), pp. 347–359. https://www.jstor.org/stable/1821637
56. De Soto, H. (2000). *The mystery of capital: Why capitalism triumphs in the West and fails everywhere else*. Basic books.
57. Ding, S., Guariglia, A., & Harris, R. (2016). The determinants of productivity in Chinese large and medium-sized industrial firms, 1998–2007. *Journal of Productivity Analysis*, 45, pp. 131–155. https://doi.org/10.1007/s11123-015-0460-0

58. Dusha, E. (2019). Persistent inequality, corruption, and factor productivity. *The BE Journal of Macroeconomics*, 19(2), 20180021. https://doi.org/10.1515/bejm-2018-0021
59. Ellis, F. (1993). *Peasant economics: Farm households and agrarian development* (2nd Ed.). Cambridge: Cambridge University Press.
60. Engel, C., & Moffatt, P. G. (2014). dhreg, xtdhreg, and bootdhreg: Commands to implement double-hurdle regression. *The Stata Journal*, *14*(4), 778-797.
61. Engerman, S. L. and Sokoloff. K. L. (2000) 'History lessons: institutions, factor endowments, and paths of development in the new world'. *Journal of Economic Perspectives*, 14(3), pp. 217–232. DOI: 10.1257/jep.14.3.217
62. Espoir, D. K. & Ngepah, N., 2020. Effects of inequality on total factor productivity: exploring the technical efficiency channel in South Africa. *African Journal of Business and Economic Research*, 15 (4), 97–122. https://hdl.handle.net/10520/ejc-aa_ajber_v15_n4_a5
63. Espoir, D. K., & Ngepah, N. (2021). Income distribution and total factor productivity: a cross-country panel cointegration analysis. *International Economics and Economic Policy*, 18, pp. 661–698. https://doi.org/10.1007/s10368-021-00494-6
64. FAO (2002). *Land Tenure and Rural Development*. FAO Land Tenure Studies No. 3. https://www.fao.org/3/y4307e/y4307e00.htm
65. FAO (2010). Climate smart agriculture: policies, practices and financing for food security, adaptation, and mitigation. In *Hague Conference on Agriculture, Food Security and Climate Change*.
66. Fentie, A., & Beyene, A. D. (2019). Climate-smart agricultural practices and welfare of rural smallholders in Ethiopia: Does planting method matter?. *Land use policy*, *85*, 387-396.
67. Fort, R. (2007). Land inequality and economic growth: a dynamic panel data approach. *Agricultural Economics*, 37(2-3), pp. 159–165. https://doi.org/10.1111/j.1574-0862.2007.00262.x

68. Frankema, E. H., 2005. *The colonial origins of inequality: Exploring the causes and consequences of land distribution* (IAI Discussion Papers No. 119).
69. Frankema, E., 2008. The historical evolution of inequality in Latin America. Unpublished doctoral dissertation. University of Groningen.
70. Frick, F., & Sauer, J. (2018). Deregulation and productivity: Empirical evidence on dairy production. *American Journal of Agricultural Economics*, 100(1), pp. 354–378. https://doi.org/10.1093/ajae/aax074
71. Galor, O., & Tsiddon, D. (1997). Technological progress, mobility, and economic growth. *The American Economic Review*, 363–382. https://www.jstor.org/stable/2951350
72. García, A. (1967). *Reforma agraria y economía empresarial en América Latina*. Santiago de Chile: Editorial Universitaria S.A., Colección Imagen de América Latina, 2.
73. Gebru, G. W., Ichoku, H. E., & Phil-Eze, P. O. (2020). Determinants of smallholder farmers' adoption of adaptation strategies to climate change in Eastern Tigray National Regional State of Ethiopia. *Heliyon*, *6*(7).
74. Gingrich, C. D. & Garber, J. D., 2010. Trade liberalization's impact on agriculture in low-income countries: a comparison of El Salvador and Costa Rica. *The Journal of Developing Areas*, 43 (2). 1–17. https://www.jstor.org/stable/40376247
75. Greene, W., 2005a. Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis*, 23 (1), 7–32. https://www.jstor.org/stable/41770178
76. Greene, W., 2005b. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*. 126 (2), 269–303. https://doi.org/10.1016/j.jeconom.2004.05.003
77. Griffiths, W. E., Hill, R. C., & Judge, G. G. (1993). *Learning and practicing econometrics*. John Wiley & Sons, Inc. New York.

78. Gupta, D. K. (1990). *The Economics of Political Violence: The effect of political instability on economic growth*. New York: Praeger.
79. Gupta, G. S., & Singh, R. D. (1984). Income inequality across nations over time: how much and why. *Southern Economic Journal*, pp. 250-257. https://www.jstor.org/stable/1058336
80. Hadley, D. (2006). Patterns in technical efficiency and technical change at the farm-level in England and Wales, 1982–2002. *Journal of Agricultural Economics*, 57(1), pp. 81–100. https://doi.org/10.1111/j.1477-9552.2006.00033.x
81. Hayami, Y., 1998. The peasant in economic modernization. *In* C. K., Eicher and J. M. Staatz, eds. *International Agricultural Development*, 3rd edition. Baltimore: Johns Hopkins University Press, 300–315.
82. Hayami, Y. & Ruttan, V. W., 1970. Agricultural productivity differences among countries. *The American Economic Review*, 60 (5), 895–911. https://www.jstor.org/stable/1818289
83. Helpman, E. (Ed.). (1998). *General purpose technologies and economic growth*. Cambridge: MIT press.
84. Henderson, H., Corral, L., Simning, E. & Winters, P., 2015. Land accumulation dynamics in developing country agriculture. *The Journal of Development Studies*, 51 (6), 743–761. https://doi.org/10.1080/00220388.2014.989993
85. Higgins, D., Balint, T., Liversage, H., & Winters, P. (2018). Investigating the impacts of increased rural land tenure security: A systematic review of the evidence. *Journal of Rural Studies*, 61, pp. 34–62. https://doi.org/10.1016/j.jrurstud.2018.05.001
86. Hogg, R. V., Tanis, E. A., & Zimmerman, D. (2014). *Probability and Statistical Inference*, 192. Printice Hall, Upper Saddle River, NJ.
87. Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating Vector Autoregressions with Panel Data. *Econometrica*, 56(6), pp. 1371–1395. https://doi.org/10.2307/1913103

88. Hutchins, J., Gong, Y., & Du, X. (2023). The role of animal breeding in productivity growth: Evidence from Wisconsin dairy farms. *American Journal of Agricultural Economics*, 106(1), pp. 286–305. https://doi.org/10.1111/ajae.12374
89. INEI (2019). *Encuesta Nacional Agropecuaria 2019*. Ficha técnica. Dirección Nacional de Censos y Encuestas. https://proyectos.inei.gob.pe/iinei/srienaho/Descarga/FichaTecnica/701-Ficha.pdf
90. INEI (2022). *Encuesta Nacional Agropecuaria 2022*. Ficha técnica. Dirección Nacional de Censos y Encuestas. https://www.inei.gob.pe/media/encuestas/documentos/ena/2022/Ficha_Tecnica.pdf
91. Jaramillo, C. F., 2001. Liberalization, crisis, and change: Colombian agriculture in the 1990s. *Economic Development and Cultural Change*, 49 (4), 821–846. https://www.jstor.org/stable/10.1086/452526
92. Jondrow, J., Lovell, C. K., Materov, I. S. & Schmidt, P., 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19 (2–3), 233–238. https://doi.org/10.1016/0304-4076(82)90004-5
93. Kaufmann, D., Kraay, A. & Mastruzzi, M., 2010. *The Worldwide Governance Indicators: A Summary of Methodology, Data and Analytical Issues*. (World Bank Policy Research Working Paper No. 5430).
94. Kea, S., Li, H. & Pich, L., 2016. Technical efficiency and its determinants of rice production in Cambodia. *Economies*, 4, 22. DOI:10.3390/economies4040022
95. Khatri-Chhetri, A., Pant, A., Aggarwal, P. K., Vasireddy, V. V., & Yadav, A. (2019). Stakeholders prioritization of climate-smart agriculture interventions: Evaluation of a framework. *Agricultural systems*, *174*, 23-31.

96. Kifle, T. (2021). Climate-Smart Agricultural (CSA) practices and its implications to food security in Siyadebrina Wayu District, Ethiopia. *African Journal of Agricultural Research*, *17*(1), 92-103.

97. Kifle, T., Ayal, D. Y., & Mulugeta, M. (2022). Factors influencing farmers adoption of climate smart agriculture to respond climate variability in Siyadebrina Wayu District, Central highland of Ethiopia. *Climate Services*, *26*, 100290.

98. Koning, N., & Pinstrup-Andersen, P., eds., 2007. *Agricultural trade liberalization and the least developed countrie*s. Dordrecht: Springer.

99. Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review*, 45(1), pp. 1–28. https://www.jstor.org/stable/1811581

100. Landreth, H. (1976). *History of Economic Theory: Scope, Method, and Content*. Boston: Houghton Mifflin.

101. Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2), pp. 317–341. https://doi.org/10.1111/1467-937X.00246

102. Lloyd-Ellis, H. (2003). On the impact of inequality on productivity growth in the short and long term: A synthesis. *Canadian Public Policy/Analyse de Politiques*, S65–S86. https://www.jstor.org/stable/3552277

103. Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1), pp. 3–42. https://doi.org/10.1016/0304-3932(88)90168-7

104. Luebker, M. (2010). *Inequality, income shares and poverty: The practical meaning of Gini coefficients*. TRAVAIL Policy Brief No. 3, International Labor Office. https://www.ilo.org/wcmsp5/groups/public/@ed_protect/@protrav/@travail/documents/publication/wcms_145695.pdf

105. Mango, N., Makate, C., Tamene, L., Mponela, P., & Ndengu, G. (2018). Adoption of small-scale irrigation farming as a climate-smart agriculture practice and its influence on household income in the Chinyanja Triangle, Southern Africa. *Land*, *7*(2), 49.

106. Makate, C. (2019). Effective scaling of climate smart agriculture innovations in African smallholder agriculture: A review of approaches, policy and institutional strategy needs. *Environmental science & policy*, *96*, 37-51.

107. Maru, B., Maryo, M., & Kassa, G. (2022). Socioeconomic determinants of crop diversity in Bule Hora Woreda, Southern Ethiopia. *Heliyon*, *8*(5).

108. Mastromarco, C. & Ghosh, S., 2009. Foreign capital, human capital, and efficiency: A stochastic frontier analysis for developing countries. *World Development,* 37 (2), 489–502. https://doi.org/10.1016/j.worlddev.2008.05.009

109. McDowell, A. (2003). From the help desk: hurdle models. *The Stata Journal*, *3*(2), 178-184.

110. Meeusen, W. & van Den Broeck, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 435–444. https://www.jstor.org/stable/2525757

111. Mdingi, K., & Ho, S. Y. (2021). Literature review on income inequality and economic growth. *MethodsX*, 8, 101402. https://doi.org/10.1016/j.mex.2021.101402

112. Mirza, M. M. Q. (2003). Climate change and extreme weather events: can developing countries adapt?. *Climate policy*, *3*(3), 233-248.

113. Mizik, T. (2021). Climate-smart agriculture on small-scale farms: A systematic literature review. *Agronomy*, *11*(6), 1096.

114. Mo, P. H. (2003). Land distribution inequality and economic growth: transmission channels and effects. *Pacific Economic Review*, 8(2), pp. 171–181. https://doi.org/10.1111/j.1468-0106.2003.00218.x

115. Morais, G. & Braga, J. M., 2018. Irrigation and farm efficiency in Brazil. *Conference of the International Association of Agricultural Economists*, Vancouver, Canada. DOI: 10.22004/ag.econ.275987

116. Mulusew, A., & Mingyong, H. (2023). An empirical investigation of the dynamic linkages of land access and food security: Evidence from Ethiopia using system GMM approach. *Journal of Agriculture and Food Research*, 11, 100494. https://doi.org/10.1016/j.jafr.2023.100494

117. Murray, U., Gebremedhin, Z., Brychkova, G., & Spillane, C. (2016). Smallholder farmers and climate smart agriculture: Technology and labor-productivity constraints amongst women smallholders in Malawi. *Gender, Technology and Development*, *20*(2), 117-148.

118. Navarro-Castañeda, S., Arranz, J. M., Burguillo, M., & De Robertis, E. C. (2021). Land tenure security and agrarian investments in the Peruvian Highlands. *Land Use Policy*, 109, 105651. https://doi.org/10.1016/j.landusepol.2021.105651

119. Neves, P. C. & Silva, S. M. T., 2014. Inequality and growth: Uncovering the main conclusions from the empirics. *Journal of Development Studies*, 50 (1), 1–21. https://doi.org/10.1080/00220388.2013.841885

120. Neway, M. M., & Zegeye, M. B. (2022). Gender differences in the adoption of agricultural technology in North Shewa Zone, Amhara Regional State, Ethiopia. *Cogent Social Sciences*, *8*(1), 2069209.

121. Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6), pp. 1417–1426. https://doi.org/10.2307/1911408

122. Nkomoki, W., Bavorová, M., & Banout, J. (2018). Adoption of sustainable agricultural practices and food security threats: Effects of land tenure in Zambia. *Land Use Policy*, 78, 532–538. https://www.sciencedirect.com/science/article/abs/pii/S0264837718304265

123. Nurkse, R., 1966. *Problemas de formación de capital en los países insuficientemente desarrollados*, cuarta edición. México: Fondo de Cultura Económica.

124. O'Donnell, C. J. (2016). Using information about technologies, markets and firm behaviour to decompose a proper productivity index. *Journal of Econometrics*, 190(2), pp. 328–340. https://doi.org/10.1016/j.jeconom.2015.06.009

125. Olley, S., & Pakes, A. (1992). *The dynamics of productivity in the telecommunications equipment industry*. (NBER Working Paper No. 3977). DOI 10.3386/w3977

126. Otsuka, K., 2007. Efficiency and equity effects of land markets. *In* R. E. Evenson and P. Pingali, eds. *Handbook of Agricultural Economics*, 3. Amsterdam: North-Holland, 2671–2703.

127. Persson, T., & Tabellini, G. (1991). *Is inequality harmful for growth? Theory and evidence*. (NBER, Working Paper No. 3599). https://www.nber.org/papers/w3599

128. Rampa, A., Gadanakis, Y., & Rose, G. (2020). Land Reform in the Era of Global Warming—Can Land Reforms Help Agriculture Be Climate-Smart?. *Land*, *9*(12), 471.

129. Rada, N. E., & Fuglie, K. O. (2019). New perspectives on farm size and productivity. *Food Policy*, 84, pp. 147–152. https://doi.org/10.1016/j.foodpol.2018.03.015

130. Reyes, A., 1978. *Latifundio y poder político. La hacienda ganadera en Sucre*. Bogota: Editorial CINEP, Serie Colombia Agraria, 2.

131. Roco, L., Engler, A., Bravo-Ureta, B., & Jara-Rojas, R. (2014). Farm level adaptation decisions to face climatic change and variability: Evidence from Central Chile. *Environmental Science & Policy*, *44*, 86-96.

132. Robinson, S. (1976). A note on the U hypothesis relating income inequality and economic development. *The American Economic Review*, 66(3), pp. 437–440. https://www.jstor.org/stable/1828182

133. Roodman, D. (2009a). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1), pp. 86–136. https://doi.org/10.1177/1536867X0900900106

134. Roodman, D. (2009b). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), pp. 135–158. https://doi.org/10.1111/j.1468-0084.2008.00542.x

135. Ruttan, V. W. (2002). Productivity growth in world agriculture: sources and constraints. *Journal of Economic perspectives*, 16(4), pp. 161–184. DOI: 10.1257/089533002320951028

136. Šergo, Z., Poropat, A., & Ilak, A. (2009). Wheat Productivity and Plough Land Inequality in Rural Croatia. *Interdisciplinary Management Research*, V, pp. 781–791. https://core.ac.uk/download/pdf/6255887.pdf

137. Sertse, S. F., Khan, N. A., Shah, A. A., Liu, Y., & Naqvi, S. A. A. (2021). Farm households' perceptions and adaptation strategies to climate change risks and their determinants: Evidence from Raya Azebo district, Ethiopia. *International Journal of Disaster Risk Reduction*, *60*, 102255.

138. Setshedi, K. L., & Modirwa, S. (2020). Socio-economic characteristics influencing small-scale farmers' level of knowledge on climate-smart agriculture in mahikeng local municipality, North West province, South Africa. *South African Journal of Agricultural Extension*, *48*(2), 139-152.

139. Sickles, R. C. & Zelenyuk, V., 2019. *Measurement of productivity and efficiency.* Cambridge: Cambridge University Press.

140. Silvana, A., & Peršurić, I. (2013). Vineyard productivity and land inequality in rural Croatia. *Journal of Food, Agriculture & Environment*, 11(3&4), pp. 675–678.

141. Singirankabo, U., & Ertsen, M. (2020). Relations between land tenure security and agricultural productivity: Exploring the effect of land registration. *Land*, 9(5), 138. https://www.mdpi.com/2073-445X/9/5/138

142. Skevas, T., & Martinez-Palomares, J. C. (2023). Technology heterogeneity and sustainability efficiency: Empirical evidence from Peruvian coffee production. *European Journal of Operational Research*, 310(3), pp. 1192–1200. https://doi.org/10.1016/j.ejor.2023.03.025

143. Skevas, T., Massey, R., & Grashuis, J. (2022). Farmer adoption and intensity of use of extreme weather adaptation and mitigation strategies: evidence from a sample of Missouri farmers. *Climatic Change*, *174*(1), 18.

144. Solow, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, 39(3), pp. 312–320. https://www.jstor.org/stable/1926047

145. Sokoloff, K. L. & Engerman, S. L., 2000. History lessons: Institutions, factor endowments, and paths of development in the new world. *Journal of Economic Perspectives*. 14 (3), 217–232. DOI: 10.1257/jep.14.3.217

146. Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. New York: W. W. Norton & Company.

147. Suescún, C., 2013. La inercia de la estructura agraria en Colombia: determinantes recientes de la concentración de la tierra mediante un enfoque espacial. *Cuadernos de Economía*, 32 (61), 653–682. https://ssrn.com/abstract=2422894

148. Sunge, R. & Ngepah, N., 2020. Agricultural trade liberalization, regional trade agreements and agricultural technical efficiency in Africa. *Outlook on Agriculture*, 49 (1), 66–76. https://doi.org/10.1177/0030727019870551

149. Tadesse, B., & Ahmed, M. (2023). Impact of adoption of climate smart agricultural practices to minimize production risk in Ethiopia: a systematic review. *Journal of Agriculture and Food Research*, 100655.

150. Tambo, J. A., & Abdoulaye, T. (2012). Climate change and agricultural technology adoption: the case of drought tolerant maize in rural Nigeria. *Mitigation and Adaptation Strategies for Global Change*, *17*, 277-292.

151. Teklewold, H., Mekonnen, A., & Kohlin, G. (2019). Climate change adaptation: a study of multiple climate-smart practices in the Nile Basin of Ethiopia. *Climate and Development*, *11*(2), 180-192.

152. Thinda, K. T., Ogundeji, A. A., Belle, J. A., & Ojo, T. O. (2020). Understanding the adoption of climate change adaptation strategies among smallholder farmers: Evidence from land reform beneficiaries in South Africa. *Land Use Policy, 99*, 104858.

153. Totin, E., Segnon, A. C., Schut, M., Affognon, H., Zougmoré, R. B., Rosenstock, T., & Thornton, P. K. (2018). Institutional perspectives of climate-smart agriculture: A systematic literature review. *Sustainability, 10*(6), 1990.

154. Tsige, M., Synnevåg, G., & Aune, J. B. (2020). Gendered constraints for adopting climate-smart agriculture amongst smallholder Ethiopian women farmers. *Scientific African*, *7*, e00250.

155. UNCTAD. (2015). *The Least Developed Countries Report 2015. Transforming rural economies*. https://unctad.org/publication/least-developed-countries-report-2015

156. Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of Economic Surveys*, 26(1), pp. 98–128. https://doi.org/10.1111/j.1467-6419.2010.00631.x

157. Van Dijk, M., Morley, T., van Loon, M., Reidsma, P., Tesfaye, K., & van Ittersum, M. K. (2020). Reducing the maize yield gap in Ethiopia: Decomposition and policy simulation. *Agricultural Systems*, *183*, 102828.

158. Vasco, E., 1978. *Clientelismo y minifundio. Bases socio-económicas del poder político en un municipio minifundista*. Bogotá: Editorial CINEP, Serie Colombia Agraria, 3.

159. Vollrath, D. (2007). Land distribution and international agricultural productivity. *American Journal of Agricultural Economics*, 89(1), pp. 202–216. https://doi.org/10.1111/j.1467-8276.2007.00973.x

160. Wang, L., Chen, Y., & Ding, S. (2022). Examining the Impact of digital finance on farmer consumption inequality in China. *Sustainability*, 14(20), 13575. https://doi.org/10.3390/su142013575

161. Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3), pp. 112–114. https://doi.org/10.1016/j.econlet.2009.04.026

162. Zegarra, E. (1999). *El mercado de tierras rurales en el Perú*. CEPAL, Serie Desarrollo Productivo, No. 63. https://www.cepal.org/sites/default/files/publication/files/4667/S99117460_es.pdf

163. Zegarra, E., Escobal, J. & Aldana, U. (2008). Titling, Credit Constraints and Rental Markets in Rural Peru: Exploring Channels and Conditioned Impacts. (Inter-American Development Bank Publications, Working Paper No. 3707).

164. Zegeye, M. B. (2021). Adoption and ex-post impact of agricultural technologies on rural poverty: evidence from Amhara region, Ethiopia. *Cogent Economics & Finance*, *9*(1), 1969759.

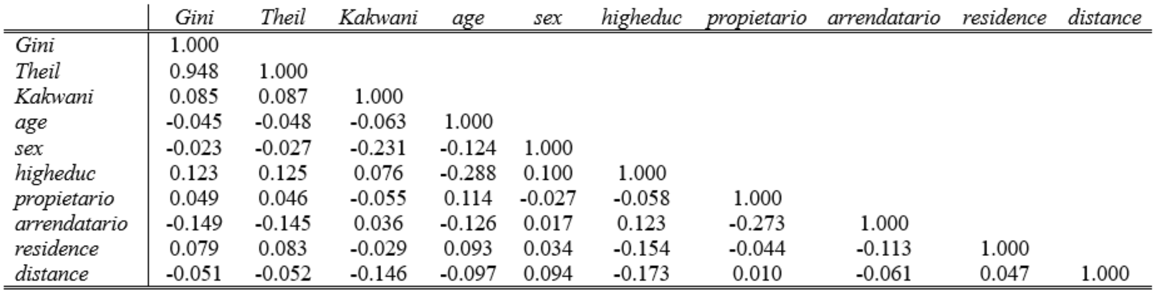
165. Zerssa, G., Feyssa, D., Kim, D. G., & Eichler-Löbermann, B. (2021). Challenges of smallholder farming in Ethiopia and opportunities by adopting climate-smart agriculture. *Agriculture*, *11*(3), 192.

APPENDIX

Table A1. Correlation matrix

	Gini	*Illiteracy*	*Tariffs*	*Irrigation*
Gini	1.000			
Illiteracy	0.262	1.000		
Tariffs	0.021	0.312	1.000	
Irrigation	-0.311	0.259	0.182	1.000

Table A2. Correlation matrix



	Gini	*Theil*	*Kakwani*	*age*	*sex*	*higheduc*	*propietario*	*arrendatario*	*residence*	*distance*
Gini	1.000									
Theil	0.948	1.000								
Kakwani	0.085	0.087	1.000							
age	-0.045	-0.048	-0.063	1.000						
sex	-0.023	-0.027	-0.231	-0.124	1.000					
higheduc	0.123	0.125	0.076	-0.288	0.100	1.000				
propietario	0.049	0.046	-0.055	0.114	-0.027	-0.058	1.000			
arrendatario	-0.149	-0.145	0.036	-0.126	0.017	0.123	-0.273	1.000		
residence	0.079	0.083	-0.029	0.093	0.034	-0.154	-0.044	-0.113	1.000	
distance	-0.051	-0.052	-0.146	-0.097	0.094	-0.173	0.010	-0.061	0.047	1.000

Table A3. Correlation matrix

	Gini	*Kakwani*	*age*	*high ed.*	*experience*	*farm inc.*	*farmsize*	*tenure sec.*	*female*	*watersource*	*extension*	*credit*	*association*	*off-farm inc*	*crop out.*
Gini	1.000														
Kakwani	0.050	1.000													
age	0.038	0.035	1.000												
high educ	-0.052	-0.040	-0.303	1.000											
experience	0.006	-0.018	0.764	-0.307	1.000										
farm income	-0.212	-0.314	-0.035	0.123	-0.052	1.000									
farm size	0.066	-0.251	0.012	-0.002	0.017	0.049	1.000								
tenure sec.	-0.039	0.136	0.120	-0.053	0.110	-0.042	-0.045	1.000							
female	0.028	0.181	0.029	-0.085	-0.023	-0.090	-0.004	0.050	1.000						
water source	-0.057	-0.011	0.045	0.019	0.029	0.095	-0.001	0.059	-0.043	1.000					
extension	-0.046	-0.057	-0.045	0.080	-0.041	0.087	-0.001	0.004	-0.039	0.138	1.000				
credit	-0.099	-0.115	-0.116	0.099	-0.094	0.215	0.019	-0.015	-0.055	0.101	0.162	1.000			
association	-0.054	-0.133	-0.049	0.056	-0.024	0.097	0.020	0.024	-0.044	0.026	0.291	0.125	1.000		
off-farm inc.	-0.026	0.028	-0.375	0.233	-0.315	-0.036	-0.017	-0.019	-0.162	-0.025	0.021	0.084	-0.014	1.000	
crop output	-0.047	-0.020	-0.021	0.061	-0.055	0.100	-0.026	0.040	-0.059	0.000	-0.058	0.020	-0.027	0.070	1.000

Table A4. Zero-truncated negative binomial regression output

	Gini	**Theil**	**Kakwani**
VARIABLES	*y*	*y*	*y*
Gini	0.066		
	(0.084)		
		0.035**	
		(0.016)	
Kakwani			0.075
			(0.080)
age	0.002	0.001	0.002
	(0.001)	(0.001)	(0.001)
high educ	0.048*	0.048*	0.048*
	(0.025)	(0.025)	(0.025)
experience	0.002	0.002	0.002
	(0.001)	(0.001)	(0.001)
farm income	7.65e-07***	7.62e-07***	7.78e-07***
	(2.74e-07)	(2.66e-07)	(2.73e-07)
farm size	-0.000	-0.000	-9.18e-05
	(0.000)	(0.000)	(0.000)
tenure security	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
female	0.018	0.018	0.014
	(0.022)	(0.022)	(0.022)
water source	0.649***	0.647***	0.648***
	(0.024)	(0.024)	(0.024)
extension	0.151***	0.151***	0.151***
	(0.039)	(0.039)	(0.039)
credit	0.092***	0.094***	0.092***
	(0.035)	(0.035)	(0.035)
association	0.010	0.011	0.013
	(0.053)	(0.053)	(0.053)
off-farm income	0.032	0.033	0.030
	(0.023)	(0.023)	(0.023)
crop output	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
alpha	1.24e-09	-20.51	1.24e-09
	(•)	(•)	(•)
LR test of alpha=0			
chibar2	0.000	0.000	0.000
Prob >= chibar2	1.000	1.000	1.000
Pseudo R2	0.075	0.075	0.075
Prob > Chi2	0.000	0.000	0.000
Log-likelihood	-6347.460	-6345.376	-6347.335
Observations	3,952	3,952	3,952

Constants omitted. Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A5. Likelihood-ratio test results

	Gini	**Theil**	**Kakwani**
One-step log-likelihood (L_{1S})	-7202.405	-7200.962	-7202.396
1st hurdle log-likelihood (L_L)	-807.107	-807.726	-805.900
2nd hurdle log-likelihood (L_T)	-6347.459	-5938.338	-6347.335
Lambda (λ)	95.677	909.796	98.321
Chi-square critical value	33.409	33.409	33.409

Null hypothesis: the one-step count model is preferable to a double-hurdle model. In all cases, λ exceeds the Chi-square critical value, rejecting the null in favor of the alternate hypothesis—i.e., the double-hurdle model is preferred.

VITA

Hernan Borrero is currently a Fulbright scholar seeking a Ph.D. degree in the Agricultural & Applied Economics program of the University of Missouri-Columbia, United States. In 2018, he completed the MSc Economics program of the University of Bristol, United Kingdom, with the Colfuturo Scholarship and the Santander Postgraduate Award; and in 2016, he graduated from the International Relations and Economics bachelor's programs of Universidad del Norte, Colombia, thanks to the Surtigas Scholarship. In 2015, he earned the Mejor Saber Pro award of the Ministry of Education for obtaining the best State exam results in the field of social sciences; and in 2014, he participated in a student exchange program at the University of Aarhus, Denmark. He served as Editor-in-chief of the *Agora Económica* journal for three years, and represented his university in the regional Karate-do tournaments, obtaining silver and bronze medals.