ABSTRACT

Title of dissertation: THE ROLE OF SOILS IN PRODUCTION: AGGREGATION, SEPARABILITY, AND YIELD DECOMPOSITION IN KENYAN AGRICULTURE

Simone Pieralli, Doctor of Philosophy, 2015

Dissertation directed by: Professor Robert G. Chambers
Department of Agricultural and Resource Economics

Agricultural production relies on soils. Increasing global population and the impact of climate change threaten the sustainability of soil for agricultural production. For these reasons, it is necessary to broaden present current methodological approaches to incorporating soil into economic analysis.

The first essay proposes a methodology to aggregate quantitative soil characteristics through the use of separability theory in a Data Envelopment Analysis framework. This yields an aggregate soil-quality measure that appropriately aggregates soil characteristics. The application is to Kenyan maize farmers.

The second essay develops a nonparametric statistical test of structural separability based on a bias correction of a central limit theorem for Data Envelopment Analysis estimators developed in Kneip et al. (2015a). The proposed nonparametric test for structural separability adapts the statistical procedures to test technology restrictions present in Kneip et al. (2015b). Monte Carlo experiments determine the size and power properties of the proposed test. An empirical analysis of Kenyan

household farmers illustrates the use of the methodology.

Global needs for higher agricultural production require understanding whether the frequently noted inverse land size-yield relationship is a true empirical regularity or an artifact of data collection methods. To examine this relationship, the third essay of this dissertation generalizes productivity decomposition methods to incorporate the quantification of a soil-productivity contribution. The generalized method decomposes a yield index into separate components attributable to (1) efficiency, (2) soil quality, (3) land size, (4) variable inputs, (5) capital inputs, and (6) output mix. Nonparametric productivity accounting methods are used to decompose the inverse land size-yield relationship in a multi-output representation of the technology without specific assumptions on returns to scale. A strongly significant inverse land size-yield relationship is present among Kenyan farmers.

THE ROLE OF SOILS IN PRODUCTION : AGGREGATION, SEPARABILITY, AND YIELD DECOMPOSITION IN KENYAN AGRICULTURE

by

Simone Pieralli

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor Robert G. Chambers, Chair/Advisor
Professor Richard E. Just
Professor Marc L. Nerlove
Professor Kenneth L. Leonard
Professor Frank J. Coale

To those who made this thesis possible..

Among others..to my inspirational grandparents, to my always greatly supportive parents, and most importantly of all to my unique companion in life.

Acknowledgments

I owe my gratitude to all the people who have made this thesis possible and because of whom my learning experience in this period of my life has been one that I will remember forever.

First and foremost I'd like to thank my mentor and advisor, Professor Dr. Robert G. Chambers for giving me an invaluable opportunity to work on challenging and extremely interesting projects over the past years. Among his intricate schedule, he has always found time for meeting and discussing with me the issues of my thesis. He has always brought me up to the door of knowledge. And let me do the most important step of knocking and entering. As we share the interests in topics of production economics, we have also participated together in conferences where I could appreciate the touch that makes distinct a great intellectual also outside the walls of a university. It has been an honor and a rewarding pleasure to work with and learn from such an extraordinary human being. He also made possible, together with my wife and time, a change in how my thoughts are translated into knowledge available and, I hope, useful to others outside me.

I would also like to thank my other thesis committee members. I would like to thank Professor Dr. Richard E. Just for sharing knowledge and personal experiences with me. I would like to thank him for showing me what is commitment, with facts. He has always been there in moments of need with an understanding eye. I would like to thank also Professor Dr. Marc L. Nerlove, with whom I met numerous times for advice in econometrics. I would like to thank also Professor Dr. Frank J.

Coale for sharing his knowledge on soils and for his commitment in representing the Graduate School in my thesis committee. Lastly but not less importantly, I would like to thank Professor Dr. Kenneth L. Leonard, who shares with me interests in development economics and without whom I would have never survived the loss of Professor Dr. Bruce L. Gardner. To Professor Dr. Bruce Gardner goes my gratitude for having shown me what is American hospitality in the very first days of my stay in the United States and having introduced me to the 'Maryland community'.

To such a school in the United States of America I could arrive only thanks to the Fulbright Graduate Scholarship sponsored by the United States Department of State and the Italian Ministry of Foreign Affairs. To both I owe my gratitude. I would like to thank Professor Dr. Bruce R. James for sharing with me his knowledge on soils and for clarifying the importance of soil in production economics.

I also owe my gratitude to Professor Dr. Martin Odening in the Department of Agricultural Economics at Humboldt University in Berlin for having believed in my capabilities and for having provided me with the possibility to research in many interesting topics and train my lecturing capabilities under his supervision in the process of finalizing my thesis. I would like to thank also Professor Dr. Silke Hüttel who has been a vivid stimulus in always fostering my knowledge. I owe gratitude as well to the administrative staff in the Department and to the facilities of the Humboldt University in Berlin, especially for its computing capabilities. I used extensively a Cluster Of Unix Machines in the Grimm-Zentrum for making the simulations of my thesis in affordable time.

I also would like to thank Professor Dr. Léopold Simar and Professor Dr. Paul

Wilson for having shared their knowledge at repeated moments in understanding statistical theory of nonparametric efficiency estimators. Many other people have made this thesis possible sharing comments or interacting with me at different moments: among them Professor Dr. Andreas Lange, Professor Dr. Erik Lichtenberg, Professor Dr. Camilla Mastromarco, Professor Dr. Angelo Zago, Professor Dr. Tim Coelli, Professor Dr. Robin Sickles, Professor Dr. Spiro Stefanou, Professor Dr. Giannis Karagiannis, Professor Dr. Alfons Oude Lansink, and Professor Dr. Rolf Färe.

I would also like to acknowledge help and support from different generations of staff members in the Department of Agricultural and Resource Economics in the University of Maryland. Great help in library searches came from Ms Katherine Faulkner and in technological issues always from Mr Jeff Cunningham: he provided great guidance into computing and saw many times my indiscriminate use of computing power but never said anything against it. Ms Liesl Koch, Ms Jane Doyle, Ms Barbara Burdick, and Ms Pin Wuan Lin have helped a lot in administrative issues.

I also thank Dr. Takashi Yamano in the National Graduate Institute for Policy Studies for the use of the data set, also the non-public soil data, which made the difference in the early stages of my research (and to Ms Maria Maratou who told me about them).

I owe my deepest thanks to my family - my grandparents, my parents, my relatives and especially my companion in life, who are always there when I need them, both for material and immaterial help. My wife, in particular, helped me in understanding the role of communication in research and provided editorial help in

revising the thesis.

My colleagues in the Department of Agricultural and Resource Economics at the University of Maryland have helped me in surpassing challenging times and in bettering my knowledge. Among others there are Dr. Geret De Piper, John Roberts, Dr. Kota Minegishi, Professor Dr. Jeff Flory, Dr. Adan Martinez, Dr. Upumanyu Datta, Dr. Tim Essam, Dr. Dennis Guignet, Dr. Michelle Brock, Professor Dr. Ariel Ortiz-Bobea, Dr. Kabir Malik, Dr. Trang Tran, Professor Dr. Eduardo Nakasone, Erin Mastrangelo, Professor Dr. Michele Baggio, Professor Dr. Marcella Veronesi, Professor Dr. Dimitrios Reppas, Daniel Voica, and Cristobal Ruiz-Tagle. My colleagues in the Department of Agricultural Economics at the Humboldt University in Berlin have also helped me in passing through difficult moments and bettering my knowledge: among others Dr. Günther Filler, Dr. Matthias Ritter, Dr. Zhiwei Shen, Dr. Rashmi Narayana, Christina Wagner, Stefan Kersting, and Dr. Sergio Villamayor-Tomas. I would like to thank Harald Fuller-Bennett and Professor Dr. Benjamin Deitchman for being friends, and a colleague and friend from the United Nations Development Programme from Egypt Hesham Mostafa for making clear to me what is development.

I would like to acknowledge financial support from the University of Maryland, the United States Department of State, and the Humboldt University in Berlin.

It is impossible to remember all, and I apologize to those I have accidentally left out.

Lastly, thank you all and thank God!

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Background

Agricultural production relies on soils. Increasing global population and the impact of climate change threaten the ability to use soil in a sustainable manner for agricultural production. For these reasons, it is necessary to broaden present methodological approaches to incorporating information on soil quality into economic measures.

Most researchers use land area as the only way to measure the land input used in agricultural production. Land, in fact, is a multidimensional factor. In addition to the area dimension, there are other soil-quality characteristics, such as carbon content, clay content, acidity, or the amount of nitrogen. These factors are critical to understanding the contribution of soil to production and productivity.

Soil characteristics have typically not been adequately considered and relegated to an unobservable error term. This treatment has concealed their contribution to agricultural production and implicitly assumed their separability from other inputs. Bhalla and Roy (1988) and Lamb (2003) considered these characteristics as omitted variables. Others (Sherlund et al., 2002; Abdulai and Binder, 2006; Di Falco and Chavas, 2009; Fuwa et al., 2007; Chang and Wen, 2011) included only qualitative soil characteristics in a regression analysis framework.

The interest in understanding the sustainability of the agricultural production process recently triggered the collection of extended data sets, which include quantitative soil characteristics. This newly available data has allowed researchers to consider soil characteristics quantitatively at the household level. Notwithstanding the availability of new data, Barrett et al. (2010) have assumed additive separability of measured soil characteristics on production outcomes. An exception is the paper by Hailu and Chambers (2012) who define, without necessarily assuming separability, a soil-quality Luenberger indicator and estimate it in a Bayesian framework.

This dissertation contributes to current knowledge with methods that identify the role of soils in production and productivity in three essays. These methodologies consist of creating a quantitative separable soil-quality measure, testing for its separability, and reconsidering the role of soil productivity in the inverse land size-yield relationship.

The first essay examines how separability can be maintained within a production technology without imposing specific assumptions about parametric representation, homotheticity, or returns to scale. Separability helps avoid an assumption of input free disposability. This is particularly important because it permits the presence of negative marginal products for high amounts of certain soil characteristics. This method builds on the observation that soil quality does not increase monotonically with increases in soil components such as, for example, soil carbon. The result of this aggregation process is a soil-quality measure that is not necessarily monotonically increasing.

The second essay develops a new statistical test for separability in a nonparametric context. The slow rate of convergence of the nonparametric estimators used requires a bias correction (Kneip et al., 2015a). Recently, Kneip et al. (2015a) have developed modified central limit theorems for nonparametric estimators of efficiency scores. Additionally, Kneip et al. (2015b) developed tests for other assumptions on the technology, such as difference of means, returns to scale, and convexity. The methods in the second essay use these developments to construct a test for separability.

The third essay develops a generalized productivity accounting framework that decomposes the contributions to yield from different groups of inputs and outputs, and quantitative soil characteristics. Soil productivity is often thought to mask heterogeneity that could cause an apparent inverse land size-yield relationship. One of the hypotheses of this essay is that soil productivity is negatively correlated with land size. If soil productivity were decreasing with size, an omitted variable bias might promote an apparent statistical negative relationship between yield and land size. The productivity accounting framework in the third essay reconsiders the inverse land size-yield relationship for a group of Kenyan household farmers.

The methods developed in this thesis are applied to a detailed survey of 452 Kenyan household farmers from 61 sublocations, made available by the National Graduate Institute for Policy Studies (21st century Center of Excellence Program) in Japan. This is one of the only Sub-Saharan African data sets that analyzes soil quantitatively at the household level. The sublocations are shown as dots in the map in figure 1.1. Although here applied to the Kenyan case, the methods developed in this thesis can be applied to any production technology. This thesis expands the literature on data envelopment by showing how to maintain the separability hypothesis without parametric assumptions on the technology (other than piecewise linearity). This method can be applied to any instance in which the technology is assumed separable. This separability assumption is often made to simplify the estimation of a technology. Methods in the second essay of this dissertation show how to test separability. In the same manner, the productivity methods developed in the third essay are also amenable to extensions to parametric measurements of non-radial distance functions.

1.2 Literature review

The three essays of this thesis relate to the literature on separability, soil quality, and on the inverse relationship between land size and yield. The first two essays adapt separability theory to the creation of a soil-quality measure. Separability has been considered for many years as a critical concept in economics. This dissertation uses a definition of separability (Blackorby et al., 1978) that requires nested upper contour sets between separable elements of functional correspondences. The first essay in this thesis operationalizes this definition in a Data Envelopment Analysis (DEA) framework.

This dissertation considers land as a multidimensional factor. For this reason, the literature on soil quality is critical to understand the multidimensionality of the land production input. Jaenicke and Lengnick (1999) have assumed multiplicative separability between soil quality variables (organic content, soil type, etc.), and other inputs and outputs. They define a soil quality index as the ratio between two distance measures: one including soil inputs and one excluding them.

Separability, in this thesis, is assumed directly between a sub-group of inputs, and other inputs and outputs. These methods do not need assuming multiplicative separability (Jaenicke and Lengnick, 1999). Moreover, differently from previous tests for separability (Woodland, 1978; Diewert and Wales, 1995), methods in this dissertation do not require assumptions on the functional form approximating the technology, apart from piecewise linearity.

Many authors (Woodland, 1978; Diewert and Wales, 1995; Banker, 1996; Emran and Alam, 1999; Berndt and Christensen, 1973; Blackorby et al., 1977) have considered testing for separability a crucial part of production economics. A revival of the interest in the separability problem has occurred recently (Simar and Wilson, 2007; Daraio et al., 2010) but these methods do not consider the bias of slowly converging nonparametric efficiency estimators (Kneip et al., 2015a). The second essay in this thesis adapts the methods of Kneip et al. (2015b,a) to the case of separability in inputs.

The third essay relates to the extensive debate on the inverse relationship between land size and productivity initiated by Chayanov (1926). Competing empirical explanations have been offered for this relationship: from the imperfection of factor markets (labor or land), to the unobservability of soil characteristics, to the presence of errors in measurement of variables. In particular, lack of soil quality or labor data (Bhalla, 1988) can cause a spurious negative relationship due to omitted variable bias. Even when soil quality and labor data are not omitted (Barrett et al., 2010), the inverse land size-productivity relationship remains difficult to explain. In Barrett et al. (2010) neither soil quality measurements nor the unobserved household specific shadow input and output prices explain the relationship.

Spillman (1915) is probably the first to study the importance of the size of the farm business. Spillman explains how in the United States the economic viability of the "little farm well tilled" was overturned by mechanization. The smallest effective area for a farm is found to be the one that gives employment to productive labor of the family. Chayanov (1926) also recognizes the presence of limiting factors (in particular, land) for a family to fully exploit the potential of workforce. But Chayanov (1926) maintains that demographic pressures lead not to fluctuations in the area cultivated (p.116) but to changes in agricultural intensity.

While already Spillman (1915) recognized that one of the most important factors in the yield of crops is the soil itself, many studies in the literature afterwards explain the inverse relationship with imperfections in the labor market. Sen (1962) rationalizes farmers' observed choices with imperfections within the labor market in Indian agriculture. He also shows that, while profitability of agriculture increases with the size of holding, productivity decreases with size.

Saini (1971) is the first to study the relationship between product and acreage, both with aggregated and with disaggregated Indian farm level data. With constant returns to scale, Saini (1971) explains the inverse size-yield relationship with higher labor input per acre in small farms. With a similar explanation, Berry and Cline (1979) estimate a negative land size-yield relationship using ordinary least squares in

India and Chile. They show that, proportionally, in large farms, presence of capital is higher than available labor while labor is definitely more present in small farms. Feder (1985) explains the inverse relationship between farm size and productivity with supervision effects over hired labor. Piette (2006) shows in Northeast of Brazil the inverse farm size-yield relationship to be caused by labor market imperfections. Recently, with a data set on Rwandan farms, Ali and Deininger (2014) also explain the negative relationship with imperfections in the same market.

One author that offers an alternative explanation for the inverse size-productivity relationship than labor and land market imperfections is Bardhan. Bardhan (1973) uses a Cobb-Douglas production function to estimate the relationship between land net sown area and value of output crop production in Indian farms.[1] Bardhan (1973) explains the inverse relationship not only with factor market imperfections but also with uncertainty in production.

Carter (1984) is the first to acknowledge that the inverse relationship could be caused by differences in mode of production between small and large farms in the state of Haryana in India. Also other studies recognized this difference in production as a potential explanation of the inverse relationship. Helfand and Levine (2004), as Piette (2006), study Brazil (Center-West) and observe that small and large farms do not usually have the same production function. Helfand and Levine (2004) use total factor productivity measures obtained from DEA and regress these scores on environmental variables, thereby obtaining a U-shaped relationship similar to Kimhi

[1]Considering uniquely value of output per acre assumes that the crop composition is similar for each farm. This does not take into account crop composition effects.

(2006).[2]

Kimhi (2006) structures sequentially the production process in two stages to explain the inverse plot size-yield relationship among Zambian maize farmers. Land allocation decisions depend on the information set available to farmers at the time of planting. Instead, a larger information set affects yield after planting. Yield is positively associated with plot size. The inverse relationship for small plots appears only when endogeneity of plot size is corrected by employing a Heckman two-stages selection model.

In his pioneering work, Barrett (1996) rationalizes the inverse land size-yield relationship with a two-period agricultural household production model. The agricultural household both produces and consumes under price uncertainty at the time labor allocations are made. This model shows that a land distribution and price risk can produce an inverse relationship between farm size and yield among Madagascaran farmers. Price risk has different effects on small and large farms. On small farms, price risk induces food security stress and causes a higher application of labor. On big net seller farms, instead, price risk reduces the use of costly inputs when revenues are stochastic. No difference is allowed in this study in cropping patterns or in village-level effects and, more importantly, among soil-quality endowments.

Different soil-quality endowments are considered as omitted variables for the first time by Bhalla and Roy (1988). They estimate the relationship with ordinary least squares including qualitative soil-quality variables in India and in 70% of the

[2]Separating scale economies from pure technical efficiency eliminates the inverse relationship with farm size among Honduran farms (Gilligan, 1998).

regions the relationship disappears. Bhalla (1988) uses soil-quality measurements of type, color, and depth at the farm level. These characteristics are regressed linearly to estimate soil quality, in addition to number of land fragments and sources of irrigation. A quality-adjusted land production function is estimated with constant elasticity of substitution. Bhalla (1988) discovers that quality-adjusted land area measures are much more egalitarian than unadjusted land area measures.

After Bhalla and Roy (1988), Benjamin (1995) considers the problem of the omission of the soil quality variables acknowledging the endogeneity of the farm and harvested-area size variables. Benjamin (1995), in a two-stages least squares regression on data from Java, uses population density, presence of a city, and number of individuals between 10 and 15 years old as instrumental variables. Despite these instrumental variables only explain 12% of the variation of the logarithm of area harvested in the sample, the inverse relationship disappears once the instruments are included. The actual relevance of these instruments and exogeneity to the outcome variable might be subject to critique.

Lamb (2003) studies the inverse relationship among farms in the data set of the International Crop Research Institute for the Semi Arid Tropics (ICRISAT). Unobserved soil quality eliminates the inverse land size-productivity relationship when profit (but not when labor demand) is the regressand. In the random effects estimates, land and labor market failures, together with soil quality, eliminate the inverse relationship when male labor (but not when female labor) is the dependent variable. The relationship is stronger under fixed effects but it is eliminated when farmed area is adjusted for the possibility of measurement error. Carletto et al.

(2013) use a newly extended data set that compares self-reported land size and size measured through global positioning system (GPS) among Ugandan farmers. However, they find that there is an even stronger negative land size-yield relationship if size is measured with GPS.

Assunçao and Braido (2007) motivate the inverse farm size-yield relationship in the ICRISAT data not with market failures affecting yield at the household level but with plot-specific unobservable variables. Nonetheless, no study until recently has considered specific soil-quality measurements, either at the household farm or at the plot level.

With a unique data set that includes plot-specific quantitative soil-quality characteristics, Barrett et al. (2010) reconsider the inverse relationship. They have precise measurements on soil-physical properties for multiple rice plots of Madagascaran farmers in 2002. The study considers jointly omitted soil-quality and factor markets' imperfections (both labor and land). However, only very little of the inverse relationship is explained by market imperfections and almost nothing by soil-quality physical characteristics. Barrett et al. (2010) estimate constant returns to scale Cobb-Douglas production functions (assuming full efficiency), conditioning linearly on soil characteristics.

Verschelde et al. (2011) estimate non-parametrically an inverse relationship among small farmers in Burundi between value of agricultural production and size. They find positive impact of soil quality but they only record perceived soil quality. Moreover, they do not consider crop composition effects by taking into account only value of crops.

The inverse land size-productivity relationship has often been analyzed (Assunçao and Braido, 2007; Bardhan, 1973; Barrett et al., 2010) assuming a Cobb-Douglas production function with constant returns to scale. However, the use of a Cobb-Douglas functional form implies a unitary elasticity of substitution that disguises different degrees of input substitutability as allocative inefficiency. In addition, the use of a production function implies all agents operate in a technically efficient manner. Instead, there are possibly many cases in which incentives are such that agents use inefficient input mixes and produce inefficient output quantities.

The inverse land size-productivity problem has usually been studied by regressing yield on land size (both in levels and in logarithmic terms) while conditioning on other factors (among which inputs and, seldom, soil-quality). If one conditions linearly on simple soil-quality characteristics, substitution possibilities among soil properties and other inputs are not considered.

The issues raised in this literature review provide a context for the following chapters that consider methodologically the role of Kenyan soils in production and productivity terms.

1.3 Figures

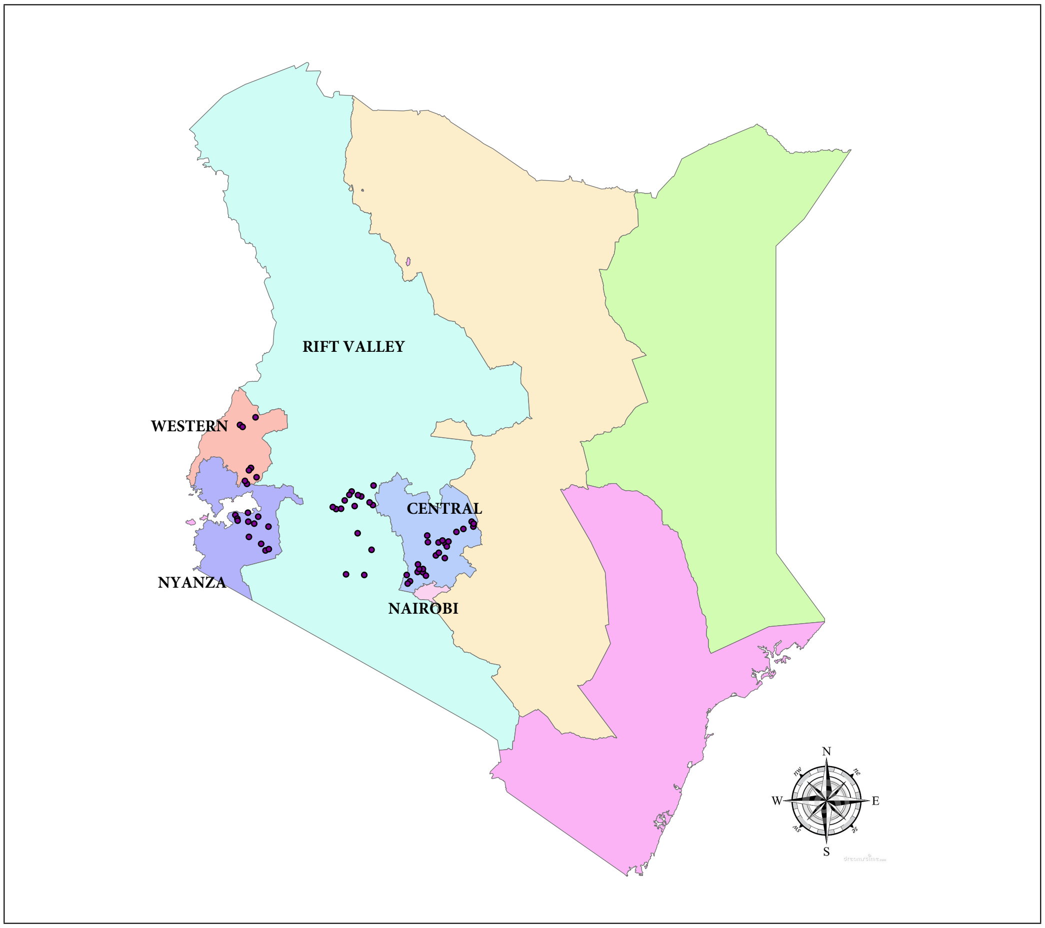


Figure 1.1: Map of sublocations in Kenya where households are surveyed

Chapter 2: Aggregation and separability in Data Envelopment Analysis

2.1 Introduction

Even though soil quality is a particularly critical factor of production, no commonly agreed measure of soil quality exists. Lacking proper soil-quality measures, a number of studies have attempted to examine the role of soil quality in production via less direct methods. For example, a series of studies have assessed the role of soil quality through the use of qualitative variables, such as slope, soil color, soil type, and soil depth (Sherlund et al., 2002; Abdulai and Binder, 2006; Di Falco and Chavas, 2009; Fuwa et al., 2007; Bellon and Taylor, 1993; Chang and Wen, 2011). Some studies (Marenya and Barrett, 2009a,b; Barrett et al., 2010) have included quantitative soil characteristics, such as soil carbon, in a regression setting. Others have used quantitative soil characteristics as freely disposable inputs to create a soil-quality index in a radial context (Jaenicke and Lengnick, 1999)[1] or a soil-quality

[1] Jaenicke and Lengnick (1999) obtain a soil-quality index multiplicatively separable from other input and output levels by imposing radial homotheticity. In another context, Färe et al. (1995) first define an intertemporal quality index without exploiting separability and then use multiplicative separability to derive a quality index only function of quality attributes.

indicator (Hailu and Chambers, 2012).

This study proposes a methodology to aggregate production factors[2] through the use of separability theory (Blackorby et al., 1978). This methodology is applied to create a soil-quality measure by aggregating quantitative soil characteristics from maize plots of Kenyan farmers. Applications of this methodology are however possible in many other instances in economics, where the measurement of a separable aggregate is appropriate.

The next section introduces the model and the following section explains the calculation methods. The fourth section describes the data set from Kenya. The fifth section presents the results and the discussion while the following section concludes.

2.2 Methods

We represent the multiple outputs by $\mathbf{y} \in \mathbb{R}_+^S$ and $\mathbf{x} \in \mathbb{R}_+^U$ denotes a vector of inputs controlled by the producer. Land area is denoted by $l \in \mathbb{R}_+$ and soil-quality characteristics are denoted by $\mathbf{c} \in \mathbb{R}_+^Q$. The technology is described by $T \subset \mathbb{R}_+^U \times \mathbb{R}_+ \times \mathbb{R}_+^Q \times \mathbb{R}_+^S$:

$$T = \{(\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \in \mathbb{R}_+^{U+1+Q+S} : (\mathbf{x}, l, \mathbf{c}) \textit{ can be used to produce } \mathbf{y}\}. \tag{2.1}$$

T satisfies:

A.1: Convexity: If $(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{y}_1) \in T$ and $(\mathbf{x}_2, l_2, \mathbf{c}_2, \mathbf{y}_2) \in T$, then $\forall \alpha \in [0, 1]$: $\alpha(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{y}_1) + (1-\alpha)(\mathbf{x}_2, l_2, \mathbf{c}_2, \mathbf{y}_2) \in T$.

[2] In particular, we consider the case of aggregating inputs, but outputs could be aggregated similarly.

A.2: Closeness of technology T.

A.3: Boundedness of output set $Y(\mathbf{x}, l, \mathbf{c}) = \{\mathbf{y} : (\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \in T\}$, $\forall(\mathbf{x}, l, \mathbf{c}) \in \mathbb{R}_+^{U+1+Q}$.

A.4: Strong disposability of outputs: if $(\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \in T$ then $0 \leqq \mathbf{y}' \leqq \mathbf{y} \Rightarrow (\mathbf{x}, l, \mathbf{c}, \mathbf{y}') \in T$.

A.5: Strong disposability of inputs $(\mathbf{x}, l)$: if $(\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \in T$ then $(\mathbf{x}', l') \geqq (\mathbf{x}, l) \Rightarrow (\mathbf{x}', l', \mathbf{c}, \mathbf{y}) \in T$

A.6: $V(\mathbf{y}; \mathbf{x}, l) = \{\mathbf{c} : (\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \in T\}$ is a convex set $\forall$ $(\mathbf{y}; \mathbf{x}, l)$.

To represent the technology, we use an output oriented Farrell measure, which is just the reciprocal of the output distance function:

$$E(\mathbf{x}, l, \mathbf{c}, \mathbf{y}) = \max\{e \in \mathbb{R}_+ \colon (\mathbf{x}, l, \mathbf{c}, e\mathbf{y}) \in T\} \tag{2.2}$$

if $\exists$ e s.t. $(\mathbf{x}, l, \mathbf{c}, e\mathbf{y}) \in T$ and 0 otherwise, and where $E : \mathbb{R}_+^U \times \mathbb{R}_+ \times \mathbb{R}_+^Q \times \mathbb{R}_+^S \to \mathbb{R}_+$. By **A.4**

$$E(\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \geq 1 \Leftrightarrow (\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \in T \tag{2.3}$$

so that $E(\mathbf{x}, l, \mathbf{c}, \mathbf{y})$ is a complete function representation of the technology.

2.2.1 Using separability to create an aggregate soil-quality measure

To create a measure of soil quality, we impose a separable structure upon $E(\mathbf{x}, l, \mathbf{c}, \mathbf{y})$. Specifically, we assume $\mathbf{c}$ are separable from $\mathbf{x}, l, \mathbf{y}$:

$$E^s(\mathbf{x}, l, g(\mathbf{c}), \mathbf{y}) = E(\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \tag{2.4}$$

where $E^s : \mathbb{R}_+^U \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+^S \to \mathbb{R}_+$ and $g : \mathbb{R}_+^Q \to \mathbb{R}_+$. $g(\mathbf{c})$ is interpretable as an aggregate of multiple soil-quality characteristics.

Our approach to measuring $g(\mathbf{c})$ is to rely on a fundamental result in separability theory (Theorem 3.2a and Corollary 3.2.0a of Blackorby et al., 1978) that demonstrates one can obtain an ordinal representation of $g(\mathbf{c})$ from the image of $E(\mathbf{x}, l, \mathbf{c}, \mathbf{y})$ through the use of reference levels of $\mathbf{x}, l,$ and $\mathbf{y}$ as follows. If the structure is truly separable, then for an arbitrary reference level $\bar{\mathbf{x}}, \bar{l}, \bar{\mathbf{y}}$:

$$E^s(\bar{\mathbf{x}}, \bar{l}, g(\mathbf{c}), \bar{\mathbf{y}}) = E(\bar{\mathbf{x}}, \bar{l}, \mathbf{c}, \bar{\mathbf{y}}) \tag{2.5}$$

which can be rewritten as[3]

$$m(g(\mathbf{c})) := E(\bar{\mathbf{x}}, \bar{l}, \mathbf{c}, \bar{\mathbf{y}}) \tag{2.6}$$

so that we can recognize $E(\bar{\mathbf{x}}, \bar{l}, \mathbf{c}, \bar{\mathbf{y}})$ as an ordinal soil-quality measure. Our aggregate $m(g(\mathbf{c}))$ quantifies, in output terms, the product of given soil characteristics for specific reference levels of other inputs and outputs. If the technology is truly separable, the change in reference levels only shifts the value of the measure but not its ordering.

This measure is different from the measures in Jaenicke and Lengnick (1999) and Färe et al. (1995). Our aggregate measure is equivalent to theirs, with appropriate disposability assumptions, if our measure is independent of the level of inputs and outputs, i.e. if the technology is homothetic in outputs and in inputs to be

[3]My thanks for showing me the possibility of writing this go to Professor Dr. Robert G. Chambers.

aggregated. However, our measure maintains separability without imposing radial homotheticity.

Notice moreover that, so long as E^s is assumed to be strictly monotonic in $g(\mathbf{c})$, this structure can accommodate technologies that do not exhibit free disposability in $\mathbf{c}$ by allowing g to be non-monotonic in $\mathbf{c}$.

2.3 Calculation

We use a nonparametric DEA approximation to the technology T:

$$T_{DEA} = \{(\mathbf{x}, l, \mathbf{c}, \mathbf{y}) \in \mathbb{R}_+^{U+1+Q+S} : x_u \geq \sum_{i=1}^{n} \gamma_i x_{iu},\, u = 1, \ldots, U;$$

$$l \geq \sum_{i=1}^{n} \gamma_i l_i;$$

$$c_q = \sum_{i=1}^{n} \gamma_i c_{iq},\, q = 1, \ldots, Q;$$

$$y_s \leq \sum_{i=1}^{n} \gamma_i y_{is},\, s = 1, \ldots, S;$$

$$\text{for}(\gamma_1, \ldots, \gamma_n) \text{ s.t.} \sum_{i=1}^{n} \gamma_i = 1, \gamma_i \geq 0, i = 1, \ldots, n\} \tag{2.7}$$

where i indexes decision-making units. In particular, notice that the set of constraints on $\mathbf{c}$ allows the aggregator $g(\mathbf{c})$ to be non-monotonic in $\mathbf{c}$. The DEA approximation of the function we use to represent the technology E can be calculated as follows:

$$E_{DEA}(\mathbf{x}, l, \mathbf{c}, \mathbf{y}) = \max e \in \mathbb{R}_+$$

$$\text{s.t. } ey_s \leq \sum_{i=1}^{n} \lambda_i y_{is},\ s = 1, \ldots, S;$$

$$x_u \geq \sum_{i=1}^{n} \lambda_i x_{iu},\ u = 1, \ldots, U;$$

$$l \geq \sum_{i=1}^{n} \lambda_i l_i;$$

$$c_q = \sum_{i=1}^{n} \lambda_i c_{iq},\ q = 1, \ldots, Q;$$

$$\text{for}(\lambda_1, \ldots, \lambda_n) \text{ s.t.} \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0, i = 1, \ldots, n. \tag{2.8}$$

Our approach to measuring the soil-quality ordinal measure $m(g(\mathbf{c})) := E(\bar{\mathbf{x}}, \bar{l}, \mathbf{c}, \bar{\mathbf{y}})$ relies on a DEA empirical approximation:

$$m_{DEA}(g(\mathbf{c})) = E_{DEA}(\bar{\mathbf{x}}, \bar{l}, \mathbf{c}, \bar{\mathbf{y}}) = \max \epsilon \in \mathbb{R}_+$$

$$\text{s.t. } \epsilon \bar{y}_s \leq \sum_{i=1}^{n} \mu_i y_{is},\ s = 1, \ldots, S;$$

$$\bar{x}_u \geq \sum_{i=1}^{n} \mu_i x_{iu},\ u = 1, \ldots, U;$$

$$\bar{l} \geq \sum_{i=1}^{n} \mu_i l_i;$$

$$c_q = \sum_{i=1}^{n} \mu_i c_{iq},\ q = 1, \ldots, Q;$$

$$\text{for}(\mu_1, \ldots, \mu_n) \text{ s.t.} \sum_{i=1}^{n} \mu_i = 1, \mu_i \geq 0, i = 1, \ldots, n \tag{2.9}$$

where $\bar{\mathbf{x}}, \bar{l}, \bar{\mathbf{y}}$ denote fixed reference levels. If the technology is truly separable, different reference levels $\bar{\mathbf{x}}, \bar{l}, \bar{\mathbf{y}}$ define different ordinal transformations of $m(g(\mathbf{c}))$.

The separable version of the technology representation on the left side of (2.4) is then easily obtained as:

$$E^S_{DEA}(\mathbf{x}, l, m_{DEA}(g(\mathbf{c})), \mathbf{y}) = \max \delta \in \mathbb{R}_+$$

$$\text{s.t. } \delta y_s \leq \sum_{i=1}^{n} \eta_i y_{is},\ s = 1, \ldots, S;$$

$$x_u \geq \sum_{i=1}^{n} \eta_i x_{iu},\ u = 1, \ldots, U;$$

$$l \geq \sum_{i=1}^{n} \eta_i l_i;$$

$$m_{DEA}(g(\mathbf{c})) \geq \sum_{i=1}^{n} \eta_i m_{DEA}(g(\mathbf{c}))_i,$$

$$\text{for}(\eta_1, \ldots, \eta_n) \text{ s.t.} \sum_{i=1}^{n} \eta_i = 1, \eta_i \geq 0, i = 1, \ldots, n. \tag{2.10}$$

The rest of this section is divided in three subsections. In the first subsection a simulated example shows the present methodology. In the second subsection tests for statistical significance of the soil-quality adjustment are described. Finally, in the third subsection a possible test for the validity of the separability condition (2.4) is explained.

2.3.1 A simulated example

To illustrate our method, we propose a simulated example. Consider 10 decision making units with one input x, two outputs y_1 and y_2, and two separable inputs c_1 and c_2. Assume, for simplicity, that all units have input x at the same level of 1. Output y_1 varies from 1 to 8 and output y_2 varies from 6 to 13. We simulate output levels such that potential congestive effects are present for both separable inputs.[4] Input and output data for our example are summarized in table 2.1.

We report the results of our methodology under variable returns to scale, for maximal reference levels of input $\bar{x}$ and outputs $\bar{\mathbf{y}}$ in the second last column of table 2.1. In this column we observe that highest estimates of our aggregate measure correspond to medium values of our separable inputs.

To visualize the result we draw in figure 2.1 an interpolated surface through the simulated aggregate measure (the actual estimates are the black dots). The contour plot of this surface is in figure 2.2. This figure shows pictorially that our

[4] In particular, congestion here refers to regions of the isoquants that show negative marginal products. In other words, these are regions of the isoquants where it is possible, keeping the rest constant, to increase output by reducing one of the aggregated inputs.

aggregate measure, given the simulated input and output levels, correctly slopes up for very low separable input levels, then peaks around medium values, and finally decreases for higher values. In other words, our method is able to account correctly for congestive effects of both aggregated production factors.

Our measure can recover the quality measures obtained in Jaenicke and Lengnick (1999) and Färe et al. (1995) if homotheticity of outputs and aggregate inputs is imposed, for example as shown in Primont and Primont (1994) and recently in Olesen (2014). However this is not necessarily the case in a multi-output case for various reference levels because our measure depends on the input and output levels of reference. To illustrate how the measure varies with different reference levels, we show the results for median reference levels in the last column of table 2.1. The numeric values of the measure are different but the ordering of the units in these results is the same as at maximal reference levels.

2.3.2 Testing for significance of the soil-quality adjustment

Even though soil-quality adjustment is important from a theoretical point of view, its statistical significance needs to be tested to motivate the adjustment. The following test is proposed:

$$\text{Test} \begin{cases} H_0 : & E^s(\mathbf{x}, l, m(g(\mathbf{c})), \mathbf{y}) = E^o(\mathbf{x}, l, \mathbf{y}) \\ \text{versus } H_1 : & E^s(\mathbf{x}, l, m(g(\mathbf{c})), \mathbf{y}) \neq E^o(\mathbf{x}, l, \mathbf{y}) \end{cases} \tag{2.11}$$

where E^s is the separable version of the Farrell output measure described in the methodological section and E^o is a similar Farrell output measure where soil-quality characteristics have been omitted.

Four different methods of testing are introduced and applied to our data set. The first two methods are based on a Kolmogorov-Smirnov two-sample test of equality of distributions. Its original derivation is due to Kolmogorov (1941). The null hypothesis is that the two tested samples are from the same distribution. The test determines whether the biggest difference among empirical distribution functions is big enough to show significant deviation through the following formula: $sup \mid F(E^s) - F(E^o) \mid$. $F(E^s)$ is the empirical distribution function of the sample of separable Farrell measures E^s. $F(E^o)$ is the empirical distribution function of the sample of Farrell measures where soil-quality characteristics are omitted. The first test statistic based on a Kolmogorov-Smirnov test in this essay includes all units. The second test statistic excludes units with Farrell measure equal to 1 in both samples.[5]

The Kolmogorov-Smirnov tests proposed assume independence of the sample observations. This is not, in general, true for a distribution of Farrell measures from a finite sample. A complicated unknown correlation is present among estimated Farrell measures in finite samples (Simar and Zelenyuk, 2006). Moreover, in finite samples, Farrell measures have a probability mass around the value of 1. Kernel density estimation, which is used (Li, 1996) to test equality of distributions, is not consistent at the boundary of the support (Simar and Zelenyuk, 2006). In particular,

[5] It is to be noted that there will be more units with Farrell measure equal to 1 under $F(E^s)$ than under $F(E^o)$. So only the observations that have Farrell measure equal to 1 under $F(E^o)$, which in turn will have a Farrell measure equal to 1 also under $F(E^s)$, are deleted.

Simar and Zelenyuk (2006) propose two methods to test equality of distributions of Farrell measures by reviewing and adapting the flexible nonparametric test by Li (1996) to the presence of a probability mass around 1 and to the presence of finite sample dependence. The first method is to truncate the distribution by eliminating the units with Farrell measure equal to 1 in both samples.[6] The second method is instead to smooth the distribution of the Farrell measures.[7]

The tests by Simar and Zelenyuk (2006) are performed here by bootstrapping 2000 times the updated version of Li's test (Li et al., 2009), from either the truncated or the smoothed distributions. All of the four tests presented are nonparametric and do not require assumptions on functional form. These tests are repeated for different returns to scale assumptions.

---

[6]The units that numerically have Farrell measures equal to 1 in both samples are the units that have Farrell measures equal to 1 in E^o.

[7]Following Simar and Zelenyuk (2006), in the adjusted case, under variable and non-increasing returns to scale, we smooth the distribution of efficiency scores by adding an error term which is distributed as $U(0, min\{n^{2/(U+1+Q+1+1)}, a-1\})$, where a is the β^{th} quantile of the empirical distribution after disregarding units with measure equal to 1. β is considered here, as in Simar and Zelenyuk (2006), to be 0.05. Under constant returns to scale, in the adjusted case, the error term is distributed as $U(0, min\{n^{2/(U+1+Q+1)}, a-1\})$. In the unadjusted case, instead, errors are distributed respectively as $U(0, min\{n^{2/(U+1+Q+1)}, a-1\})$ under variable and non-increasing returns to scale, and as $U(0, min\{n^{2/(U+1+Q)}, a-1\})$ under constant returns.

2.3.3 A proposed homogeneous bootstrapping test for separability

A methodology for testing if the separability assumption in equation (2.4) holds is proposed. One way to test the separability assumption proceeds as follows:

$$\text{Test} \begin{cases} H_0: & \mathbf{c} \text{ are separable} \\ \text{versus } H_1: & \mathbf{c} \text{ are not separable.} \end{cases} \tag{2.12}$$

It is possible to construct a test ratio (TS) to test the null hypothesis H_0 against the alternative H_1 (Simar and Wilson, 2002) as:

$$TS = \frac{E^s(\mathbf{x}, l, m(g(\mathbf{c})), \mathbf{y})}{E(\mathbf{x}, l, \mathbf{c}, \mathbf{y})}. \tag{2.13}$$

At the denominator $E(\mathbf{x}, l, \mathbf{c}, \mathbf{y})$ is the non-separable version of the technology. At the numerator, instead, $E^s(\mathbf{x}, l, m(g(\mathbf{c})), \mathbf{y})$ is the separable version. Both measures are explained in the methodology.

In other words the test is

$$\text{Test} \begin{cases} H_0: & TS = 1 \\ \text{versus } H_1: & TS \neq 1. \end{cases} \tag{2.14}$$

No theoretical motivation explains whether TS should be greater than or lower than 1. The null hypothesis of separability can be rejected if the test is statistically significantly different than unity.

Following Simar and Wilson (2002),[8] sample analogs of this test ratio can be estimated via bootstrap in different ways. One test statistic can be defined as the

[8] In more recent articles other tests have been used with heterogeneous (nonparametric) bootstrapping techniques (Schubert and Simar, 2010) or generalized jackknife statistics to correct for

mean of the ratio of the bootstrapped values of the numerator and denominator:

$$TS1 = \sum_{i=1}^{n} \left\{ \frac{E^s_{DEA}(\mathbf{x}_i, l_i, m_{DEA}(g(\mathbf{c}_i)), \mathbf{y}_i)}{E_{DEA}(\mathbf{x}_i, l_i, \mathbf{c}_i, \mathbf{y}_i)} \right\} /n \tag{2.15}$$

or as the ratio of the means of the bootstrapped values of the numerator and denominator:

$$TS2 = \frac{\sum_{i=1}^{n} E^s_{DEA}(\mathbf{x}_i, l_i, m_{DEA}(g(\mathbf{c}_i)), \mathbf{y}_i)}{\sum_{i=1}^{n} E_{DEA}(\mathbf{x}_i, l_i, \mathbf{c}_i, \mathbf{y}_i)}. \tag{2.16}$$

These ratios (Simar and Wilson, 2002) represent, respectively, the average distance between the separable and non-separable frontier and the distance between the average estimated separable frontier and the average non-separable frontier. If the distance is small on average, in a statistical sense, the null hypothesis of separability shall not be rejected.

Other possible statistics include the median of ratios, the ratio of medians, the 10% trimmed mean of ratios, or the ratio of the 10% trimmed means:

$$TS3 = Med \left\{ \frac{E^s_{DEA}(\mathbf{x}_i, l_i, m_{DEA}(g(\mathbf{c}_i)), \mathbf{y}_i)}{E_{DEA}(\mathbf{x}_i, l_i, \mathbf{c}_i, \mathbf{y}_i)} \right\}_{i=1}^{n}, \tag{2.17}$$

$$TS4 = \frac{Med\{E^s_{DEA}(\mathbf{x}_i, l_i, m_{DEA}(g(\mathbf{c}_i)), \mathbf{y}_i)\}_{i=1}^{n}}{Med\{E_{DEA}(\mathbf{x}_i, l_i, \mathbf{c}_i, \mathbf{y}_i)\}_{i=1}^{n}}, \tag{2.18}$$

$$TS5 = Trim_{10} \left\{ \frac{E^s_{DEA}(\mathbf{x}_i, l_i, m_{DEA}(g(\mathbf{c}_i)), \mathbf{y}_i)}{E_{DEA}(\mathbf{x}_i, l_i, \mathbf{c}_i, \mathbf{y}_i)} \right\}_{i=1}^{n}, \tag{2.19}$$

or

$$TS6 = \frac{Trim_{10}\{E^s_{DEA}(\mathbf{x}_i, l_i, m_{DEA}(g(\mathbf{c}_i)), \mathbf{y}_i)\}_{i=1}^{n}}{Trim_{10}\{E_{DEA}(\mathbf{x}_i, l_i, \mathbf{c}_i, \mathbf{y}_i)\}_{i=1}^{n}}. \tag{2.20}$$

bias (Kneip et al., 2015a). The adaptation of these more general methodologies of testing is a contribution that involves testing size and power properties of the new tests and that is pursued in the second essay of this dissertation.

These last four statistics use robust measures of location instead of the arithmetic mean as in $TS1$ and $TS2$. Tests $TS3$ to $TS6$ are more robust statistics when underlying distributions are skewed (Simar and Wilson, 2002).

2.4 Data

The methodology is applied to a random sample of households from 61 sublocations in Kenya. The survey is named "Research on Poverty, Environment and Agricultural Technologies (REPEAT): Panel studies in Africa". Survey data were obtained from the National Graduate Institute for Policy Studies (21st century Center of Excellence Program) in Japan. Data are from early 2007 and refer to preceding short and long growing seasons. The present cross-section sample is composed of 590 households.[9] Of these, data on soil characteristics are available only for 452 families.[10] Variables used in this analysis are on 15 inputs, 4 outputs, and 2 soil-quality

[9]The households considered here are household farmers so the two terms are used interchangeably in the following to mean a household that farms. More households are included in the original sampling scheme but are left out of the analysis because they have incomplete or erroneous entries.

[10]No significant bias seems to be introduced in this sample reduction step in the output variables. Marginal effects on the probability (in a Probit regression) of being present in the final sample are insignificant for the outputs. In the case of the inputs, the resulting sample might suffer some bias: farmers with lower number of cows are more probable to be observed in the final sample. Farmers who spend more in hired labor, and use more organic, and less inorganic, fertilizers are observed more often. These apparent biases are not critical for the methodological results of this paper. They should however be considered when evaluating the representativity of our Kenyan sample.

characteristics.[11] Main agricultural output considered is harvested dry maize which is totally rain-fed. All present farmers produce maize.

The measured inputs used in maize production are seeds, land area, organic and inorganic fertilizers, family[12] worked hours, cost of temporary hired workers, and hours worked by permanent and shared workers.

As diversification options other representative outputs are harvested green maize, milk, and off-farm income. Other measured inputs available for household production are number of hand hoes, ploughs, sickles, spray pumps, livestock permanent laborers, number of milking cows, household members[13] available for work

---

[11]Data have undergone cleaning for detection of outliers, following Wilson (1993), to help diminish at its minimum problems deriving from measurement error.

[12]Faithful to the human capital approach pioneered by Jorgenson and Griliches (1967), labor input is adjusted for differences in quality, both due to education and age. To adjust labor for education, estimates of the impact of education on agricultural productivity in Kenya from the analysis by Husbands et al. (1996) are used. Given an 85% probability that a primary school-completer household head increases household profitability by 40%, we divide the expected return equally among years of primary education (8 years) to obtain a 4.25% average increase in productivity per additional year of any level of education. It is possible that different methods of partitioning the increase in productivity would have achieved a different result. But this topic is not the focus of the present contribution. The same is done for households with a primary school completer, who is not the household head. The total increase is discounted to be only 29.75% in this case. To adjust labor for age, hours worked by children who, by United Nations Children Fund classification, are considered those less than 15 years old, are divided in half.

[13]All of the family members living at the household that are above three years old are included. Most kids start completing tasks in the family at this age. Number of members present in the household is also weighted for quality in the same way as the number of worked hours in the

in the family, and family members engaged in off-farm activities.[14]

Table 2.2 shows input and output summary statistics. Agriculture is mainly manual. Median[15] household annual off-farm earnings from all sources are 41100 Kenyan Shillings ($587 February 2007 dollars). For the same time period, households in the lowest quartile earned the equivalent of about $163. The median household head education level is primary. Median applications of inorganic and organic fertilizers on maize fields are 22 kgs and 300 kgs respectively. Median household maize harvests are 540 kgs of dry maize and 170 kgs of green maize per family. Median daily time spent by a household on the maize plots is little more than 45 minutes per day. Hand hoes are common (the median is 4 hand hoes per household), while ploughs, spray pumps, and sickles are not widespread (more than half of the households do not use any). Median household size is 4.5 people, who live at home, plus another family member who works off-farm. Half of the households have 1 cow or more.

2.4.1 Soil-quality characteristics

Data on soil-physical characteristics for the largest maize plot of each household are available for mid-2003.[16] To visualize the method, our analysis focuses

previous footnote.

[14]The analysis is done at the household and not at the farm level to account for off-farm income. Because off-farm work decisions and on-farm decisions are made jointly by the same decision makers, any representation of the technology must incorporate both sources of income.

[15]Means can be seen in table 2.2.

[16]Both soil carbon and soil clay content are immutable for long periods unless lands are converted to agriculture immediately before analyzing the soils. Our input-output data are for 2005-2006,

mainly on two critical measures: one of soil structure (carbon content) and one of soil texture (clay content). Table 2.2 shows summary statistics.

Soil carbon content (in percentage points of soil weight) is a generally accepted indicator of long-term soil structure (Brady and Weil, 2008) because it varies as a result of human action but only in the long-run. An increasing number of studies have assumed that more soil carbon content leads to better soil quality because it increases the amount of nutrients available to the plants and increases infiltration and water-holding capacity. However, soil science indicates that matters are not so simple. There is no 'ideal' amount of soil carbon (Brady and Weil, 2008). A percentage of soil carbon higher than what the soil-plant-climate is able to bear can interfere with the nitrogen cycle (Brady and Weil, 2008) and be actually indicative of poor drainage of soils. Hence, in our analysis, we allow for the possibility that negative marginal effects (input congestion) may arise from too high levels of carbon content.

Soil clay content (in percentage points of soil weight) provides an important indicator of soil texture. Finer soil clay particles magnify nutrients' retention in the soil (especially nitrogen) for better plant growth, given higher cation exchange capacity of finer textured soil particles. However, negative effects can occur for high amounts of clay due to the behavior that different types of clays have with respect to water and nutrients.

Most of the households in this sample live on kaolinitic or montmorillonitic

which is close enough to 2003 to assume that significant changes in soil carbon and clay have not occurred. We assume that all maize plots have, in each household, same soil quality, on average.

clays. Kaolinitic clays (1:1-type silicate clays) do not expand depending on water and have limited water holding capacity and nutrients' adsorption. Soils high in kaolinitic clays, without considering nutrient management, might show possible congestion due to limited water holding capacity and lower level of nutrients' adsorption. Montmorillonitic clays (2:1-type silicate clays) instead expand and contract more and hold much more water. The presence of this type of clays measures closely the water holding capacity of the soil.[17] However, in particularly rainy situations, abundance of montmorillonitic clays can have negative effects on cultivation due to the extreme stickiness of soils. Therefore, our empirical analysis allows also for the possibility of congestion due to clay.[18]

While soil carbon and soil clay are two fundamental soil characteristics used in the present study, our approach can accommodate more characteristics as aggregated factors.

2.5 Results and discussion

The results of the application of our methodology to the Kenyan data appear in a surface interpolated through the estimates of our soil-quality measure (figure

[17]Water holding capacity is also linked to the distribution of pores in the soil: big pores with less total pore space (more coarse textured soils) indicate a more sandy soil (more prone to percolation of rain), while a balanced distribution of pores indicates finer textured soils with higher water holding capacity (Brady and Weil, 2008), better suited for agriculture.

[18]I would like to thank Professor Dr. Tim Coelli and Professor Dr. Angelo Zago for observing that soil characteristics are not necessarily monotonically related to quality during European Workshop on Efficiency and Productivity Analysis 2009.

2.3: actual estimates of soil quality are the black dots). Higher soil quality is obtained from values of soil carbon between 2% and 4% depending on soil clay as well. Values of soil carbon between 2% and 4% are precisely in the range in which soil scientists observe best yields. Best yields are usually observed starting from 2% of soil organic carbon (Loveland and Webb, 2003). Figure 2.3 represents our estimates under variable returns to scale (reference levels $\bar{\mathbf{x}}, \bar{l}, \bar{\mathbf{y}}$ are fixed at the maximal level).

The levels of our soil-quality measure are visible in a contour plot of this interpolated surface (figure 2.4). In figure 2.4 two dashed lines cross the empirical isoquants where one of the marginal products changes sign. These lines, called 'ridge lines' (Ferguson, 2008), delimit a region to the south-west of the peak region. This region is the substitution region where both marginal products of soil carbon and soil clay are positive. The complementary portion of the isoquants is a region where at least one marginal product is negative. In other terms, at least one soil characteristic shows congestion.

Few observations present negative marginal effects in both soil clay and soil carbon.[19] Congestive effects are more often in only one soil characteristic. Figure 2.4 shows that soil carbon, for low levels of soil clay, has negative marginal effects from 4% onwards. In other words, a decrease of soil carbon above this level would increase soil quality, given a relatively low soil clay percentage. On the other hand, even though households live in regions with clays (montmorillonitic and kaolinitic)

[19]In particular, negative marginal effects for both soil clay and soil carbon are present for high soil clay levels and medium to low carbon levels.

with potentially negative marginal effects, soil clay shows, for low levels of soil carbon, a certain congestion only after 35-40 %. The presence of negative marginal effects for clay is justified because households in this sample could have received particularly negative climatic conditions during the study period.

The same soil-quality measure can be calculated for different reference levels $\bar{\mathbf{x}}, \bar{l}, \bar{\mathbf{y}}$. Moreover, different assumptions on the returns to scale of the technology obtain different soil-quality measures. Table 2.3 is composed of the measures under variable, non-increasing, and constant returns to scale, both at the maximal and at the median reference level. For each reference level, and each returns to scale assumption, we summarize the results of our 452 estimates by reporting the mean, the standard deviation, and the decile levels. For each reference level, the mean of our soil-quality measure increases when assuming non-increasing and constant returns. On the contrary, its standard deviation decreases at each reference level when assuming non-increasing and constant returns.[20]

In table 2.4 we test for returns to scale of the soil-quality measure at the maximal reference level with homogeneous bootstrapping test statistics ($TS2$ to $TS6$) with same tests as in Simar and Wilson (2002).[21] The results show that returns

[20] At the maximal reference level some estimates are zero in this table. These are feasible cases where one soil characteristic is either very high or very low. Some infeasibilities are instead present at the median reference level: 92 under variable returns, 21 under non-increasing returns, and 16 under constant returns. These are reported as zeros in this table. The infeasibilities are due to the choice of the very restrictive median reference level.

[21] $TS1$ is not robust in this case because we have some estimated soil-quality measures equal to zero.

to scale globally constant cannot be rejected at 95% or at 99% confidence levels, depending on the test considered. Non-increasing returns can never be rejected at 95% confidence level against variable returns.

Once we obtain soil-quality measure estimates $m_{DEA}(g(\mathbf{c}))$, we incorporate them in the separable version of the technology (E^s) in equation (2.4). Separable Farrell output measure estimates (E^s_{DEA}) are reported in table 2.5 under variable, non-increasing, and constant returns to scale.[22] As expected, average measures are decreasing when allowing for variable returns to scale.

These separable measures can be compared to Farrell output measures, calculated either excluding soil characteristics, or including soil characteristics as non-separable inputs (in the first and second half of table 2.6, respectively). These measures in table 2.6 and their standard deviations vary in the same way as the separable measures in table 2.5. In particular, when excluding soil characteristics (in the first half of table 2.6), means are higher than in the separable version (E^s_{DEA}). Means are instead lower when including soil characteristics as non-separable inputs (in the second half of table 2.6).

In the following subsections we discuss three critical aspects of the proposed quality measure. In the first subsection we evaluate the stability of the ranks of

[22] In the cases of an infeasible aggregate measure at the median reference level we used zero to calculate the separable measure E^S_{DEA}. Substituting the missing measures at median reference level with the mean value of the available measures, which would neutralize their impact on the separable measure, did not seem to change much the distribution of our separable measure at the median level. Moreover other choices of the reference levels do not pose the same problem.

our ordinal soil-quality measure. In the second subsection we ascertain whether our quality measure makes a significant difference in evaluating the technology. In the third subsection we explore the robustness of the quality measure to the key assumption of separability.

2.5.1 Soil-quality measure robustness checks

Since we defined an ordinal soil-quality measure we are interested in the stability of its ranks. We test the presence of a monotonic correlation among sets of soil-quality measures by performing Spearman coefficients of correlation tests.

We consider the robustness of our soil-quality measure either within or across returns to scale assumption in table 2.7.[23] Ranks are stable and measures have high correlation within same returns to scale assumption. Absence of correlation among soil-quality measures calculated at different reference levels is rejected (p-value$<$0.01 in first half of table 2.7).

We also obtain stable ranks when testing across different returns to scale assumptions (p-value testing absence of correlation is less than 0.01 in second half of table 2.7). On one hand, correlation is higher between constant and non-increasing returns at the median reference level. On the other hand, correlation is instead higher between non-increasing and variable returns to scale at the maximal reference level.

[23]Tests are done excluding the missing values at the median reference level. If we substitute zero to the missing values, results are similar.

2.5.2 Significance of soil-quality adjustment

Even though adjusting for soil quality is important from a theoretical point of view, it is more difficult to ascertain its statistical relevance. We perform four tests with measures calculated at maximal reference levels of inputs and outputs. Two Kolmogorov-Smirnov tests are performed among soil-quality adjusted (E^s_{DEA}) and unadjusted (excluding soil characteristics) Farrell measures. In addition, two bootstrapped tests of equality using the updated version of Li's test (Li et al., 2009) are applied either among distributions that exclude the units with Farrell measure equal to 1 or among smoothed distributions of Farrell measures (Simar and Zelenyuk, 2006).

Results are common to both sets of measures (table 2.8). If units with Farrell measure equal to 1 are included, no significant difference is observed. Even though the introduction of an additional input varies the scores a priori, a statistically significant difference is not observed. This lack of significance shows that our methodology is a 'conservative' way of adjusting for soil quality. On the contrary, when units with Farrell measure equal to 1 are excluded, significant differences are observed between distributions under variable and constant returns to scale. This shows the importance of adjusting for soil quality, at least under variable and constant returns to scale.

2.5.3 Testing for separability

The validity of the soil-quality measure proposed hinges on the identifying separability assumption. Our tests for separability under different assumptions on returns to scale are presented in table 2.9 when reference levels of other inputs and outputs are maximal. The tests reject, at various significance levels, the null hypothesis of separability in only three of the six different tests under variable and non-increasing returns. Under constant returns only two tests reject separability.

A possible explanation of these rejections is that the homogeneity assumption of the frontier, intrinsic in this bootstrap testing procedure, forces our tests to reject a non-homogeneous, but separable, technology. Results from generalized methods might be different. The study of different possible outcomes from generalized separability testing methods is in the second essay of this dissertation.

2.6 Conclusions

The importance of the present contribution is in developing a general methodology for aggregating production factors (possibly non-freely disposable) in a DEA framework through the use of separability theory. A simple numerical example shows how our measure aggregates two inputs. We apply the methodology to the creation of a soil-quality measure by aggregating soil characteristics of maize plots of Kenyan farmers. Soil characteristics are aggregated via DEA linear programming to obtain a soil-quality measure that allows congestive effects of the inputs aggregated.

Indicators of soil structure (carbon) and soil texture (clay) are the soil characteristics used in the analysis. Aggregation of these soil characteristics in the Kenyan data set results in a soil-quality measure that attains its maximum exactly in the range of values that soil scientists associate to best soil quality. Once reached the maximum, soil quality decreases, for increasing values of soil attributes. Noticeable congestion is observed for soil carbon and clay.

Tests on returns to scale of the soil-quality measure cannot consistently reject constant returns to scale. Our soil-quality aggregate shows stability as an ordinal measure, both across different returns to scale assumptions and within returns to scale assumption for different reference levels of inputs and outputs. Soil-quality adjustment is significant under constant and variable returns to scale. The presented separability tests, at least partially, validate our soil-quality measure.

2.7 Figures and Tables

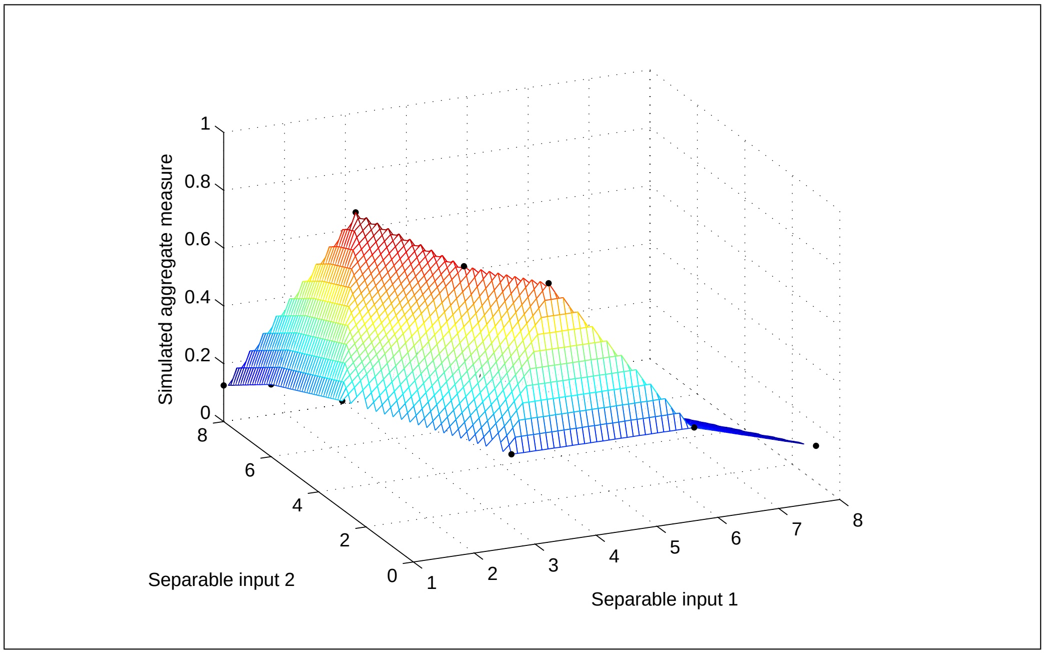


Figure 2.1: Simulated aggregate measure at maximal reference levels of other input and output for variable returns to scale

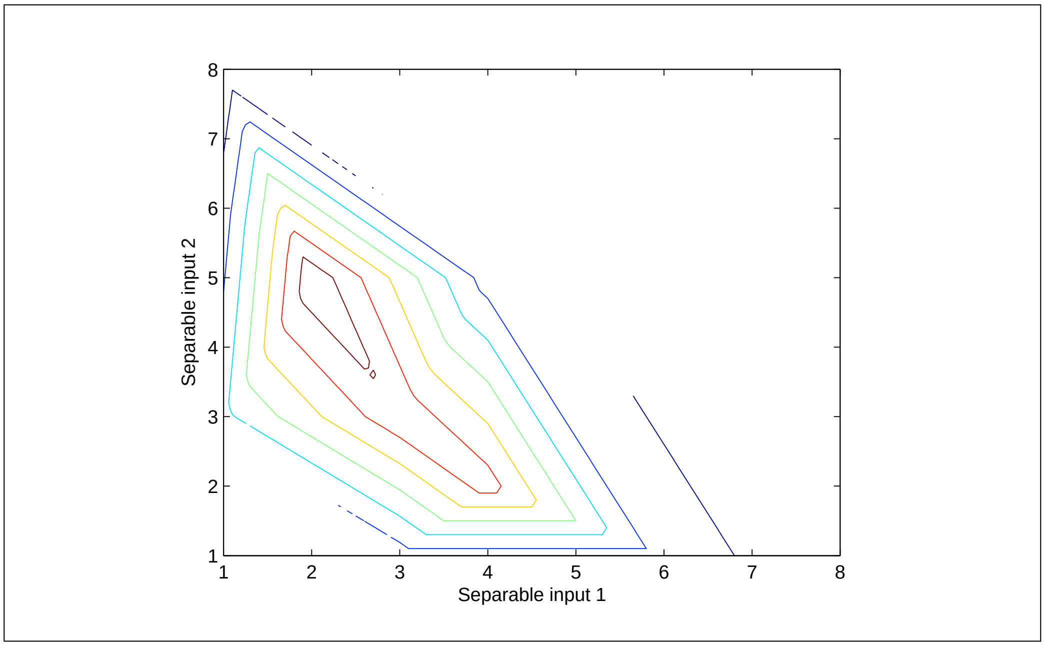


Figure 2.2: Iso-level curves of simulated aggregate measure at maximal reference levels of other input and output under variable returns to scale

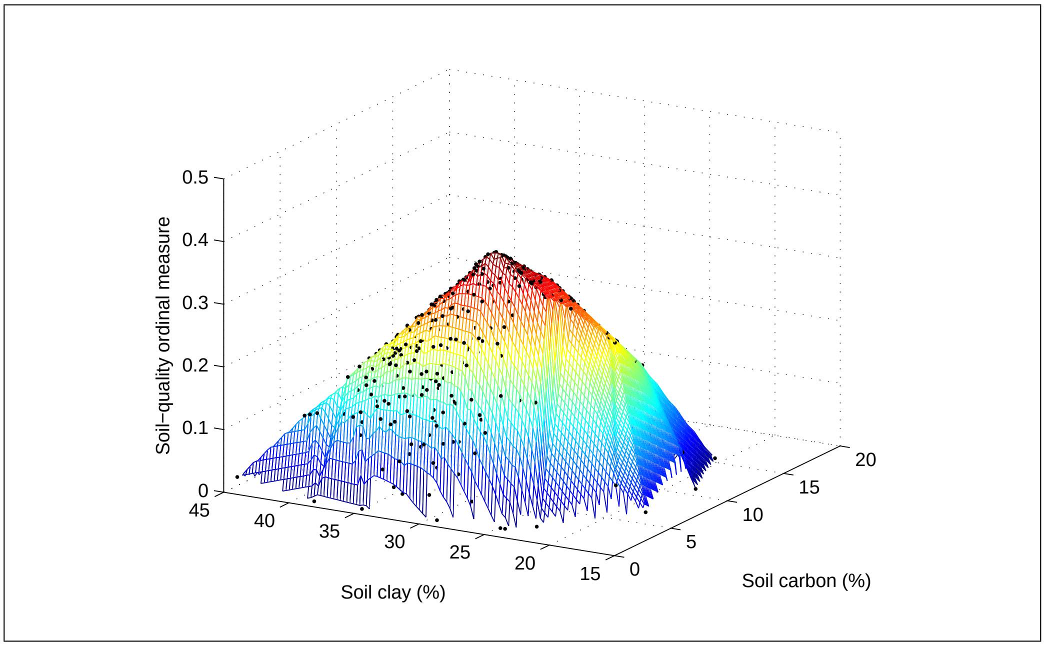


Figure 2.3: Soil-quality ordinal measure at maximal reference levels of other inputs and outputs for variable returns to scale

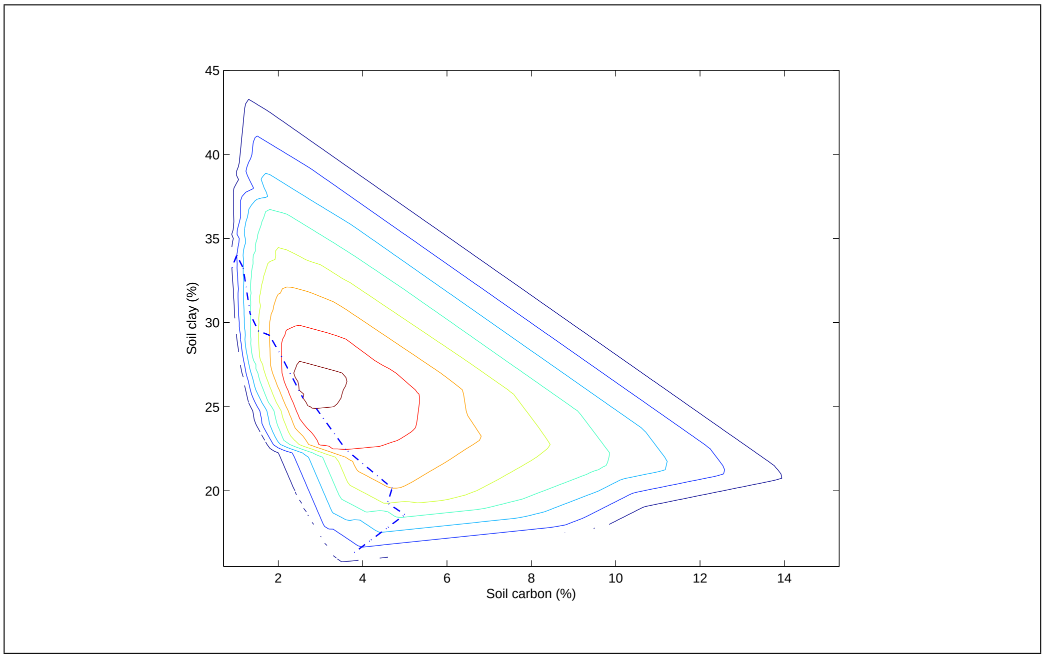


Figure 2.4: Soil-quality measure iso-level curves at maximal reference levels of other inputs and outputs under variable returns to scale

Table 2.1: Simulated inputs and outputs for 10 decision making units

Unit						Aggregate measure	
	x	y_1	y_2	c_1	c_2	VRS at max. ref.	VRS at median ref.
1	1	2	8	3	1	0.25	1
2	1	3	9	1	3	0.375	1.125
3	1	8	6	3	3	0.779	1.375
4	1	2	8	1	6	0.25	1
5	1	2	8	6	1	0.25	1
6	1	6	11	4	2	0.75	1.375
7	1	7	12	2	5	0.875	1.5
8	1	1	13	8	1	0.125	0.5
9	1	1	6	1	8	0.125	0.5
10	1	2	8	4	5	0.25	1

Table 2.2: Summary statistics of inputs, outputs, and soil characteristics

Variable	Mean	Std.Dev.	Min	Max
Inputs				
land area (acres)	1.6	1.4	0.1	14
quantity of seeds (kgs)	13.4	11	1	78
inorganic fertilizers (kgs)	46.8	74.4	0	650
organic fertilizers (kgs)	742.1	1214.6	0	9000
hired labor (cost in KSh)	2935.4	4911.9	0	48160
family labor (hours)	431.5	510.9	0	4434.8
permanent and share labor (hours)	41.7	95.1	0	963
number of hand hoes	3.9	2.2	0	15
number of ploughs	0.1	0.3	0	2
number of spray-pumps	0.4	0.6	0	2
number of sickles	0.4	0.6	0	3
livestock permanent labor	0.2	0.7	0	4.8
off-farm earners	1.1	1.1	0	6
present members	4.6	2.7	0	17.5
milking cows	1	0.9	0	5
Outputs				
total harvest dry maize (kgs)	843.7	1122.8	0	9000
total harvest green maize (kgs)	249.4	587.6	0	7180
milk (liters)	1690.9	2435.9	0	18600
off-farm income (in KSh)	85144.4	146018.3	0	1572400
Soil-quality physical characteristics				
soil carbon content (% of soil weight)	2.6	1.5	0.7	15.2
soil clay content (% of soil weight)	28.3	3.9	15.5	44.9
Observations	452			

Table 2.3: Summary statistics of ordinal soil-quality measure $m_{DEA}(g(\mathbf{c}))$ under variable, non-increasing, and constant returns to scale in correspondence of median and maximal reference vectors of other inputs and outputs

Ref. percentile	Min	10%	20 %	30 %	40 %	Median	60 %	70 %	80 %	90 %	Max	Mean	St.Dev.
Median VRS	0	0	0	1.311066	1.749283	2.031208	2.286521	2.53323	2.788758	3.009542	3.18037	1.762519	1.07262
Maximal VRS	0	0.143523	0.199804	0.24665	0.274758	0.310687	0.341032	0.360522	0.381053	0.397313	0.422828	0.288082	0.1016
Median NIRS	0	1.1728	1.634889	1.93982	2.278603	2.565739	2.753284	2.98199	3.071737	3.123764	3.180935	2.299654	0.84546
Maximal NIRS	0	0.173556	0.219499	0.254277	0.284411	0.318477	0.345194	0.367583	0.385193	0.400647	0.422828	0.30034	0.088605
Median CRS	0	1.2029	1.677857	1.960158	2.375217	2.631967	2.809173	3.017241	3.080271	3.130294	3.180935	2.347208	0.828518
Maximal CRS	0	0.176374	0.230244	0.272592	0.311685	0.348101	0.370251	0.388482	0.407405	0.430628	0.547654	0.323163	0.098814

Table 2.4: Tests for returns to scale: constant returns to scale (CRS) against variable returns to scale (VRS), and non-increasing returns to scale (NIRS) against variable returns to scale (VRS)

No. of Test	Test CRS vs VRS	90%	95%	99%
TS2	1.1218	0.8858	1.0738	2.0567
TS3	1.0359	0.9525	1.1122	2.1853
TS4	1.1204	0.8693	1.045	1.9954
TS5	1.1196	0.9184	1.1103	2.072
TS6	1.1067	0.8909	1.0799	2.0361
No. of Test	**Test NIRS vs VRS**	**90%**	**95%**	**99%**
TS2	1.0425	0.9926	1.1539	2.03
TS3	1.0012	1.0153	1.1457	2.0243
TS4	1.0251	0.9784	1.1455	2.0363
TS5	1.0362	1.0245	1.1898	2.1238
TS6	1.0336	0.9943	1.1568	2.0428

Note: The results in this table are obtained from 2000 bootstrap replications. We discarded 2 replications that presented one infeasibility in one of the 452 observations (once under variable and once under constant returns to scale).

Table 2.5: Summary statistics of separable Farrell output estimates E^{s}_{DEA} under variable, non-increasing, and constant returns to scale in correspondence of different soil-quality measures calculated for median and maximal reference percentiles of other inputs and outputs

SQ percentile	Min	10%	20 %	30 %	40 %	Median	60 %	70 %	80 %	90 %	Max	Mean	St.Dev.
Median VRS	1	1	1	1	1	1	1	1	1.212787	1.710431	6.257377	1.205156	0.547926
Maximal VRS	1	1	1	1	1	1	1	1	1.262355	1.734815	6.257377	1.212493	0.546193
Median NIRS	1	1	1	1	1	1	1	1.130205	1.413625	1.850992	6.257377	1.262616	0.597278
Maximal NIRS	1	1	1	1	1	1	1	1.14224	1.430907	1.850633	6.257377	1.270529	0.605204
Median CRS	1	1	1	1	1	1	1	1.164359	1.46146	1.855338	6.949481	1.280501	0.623518
Maximal CRS	1	1	1	1	1	1	1	1.155413	1.46146	1.852671	6.894612	1.278632	0.619042

Table 2.6: Summary statistics of conventional Farrell output measures E_{DEA} under variable, non-increasing, and constant returns to scale, excluding and including in a non-separable fashion, soil clay and soil carbon

Returns to scale	Min	10%	20 %	30 %	40 %	Median	60 %	70 %	80 %	90 %	Max	Mean	St.Dev.
Excluding													
Variable	1	1	1	1	1	1	1	1.098159	1.412962	1.849018	6.257377	1.265819	0.607498
Non-increasing	1	1	1	1	1	1	1	1.184112	1.485765	1.910524	6.257377	1.291475	0.625263
Constant	1	1	1	1	1	1	1.045105	1.247969	1.575124	2.022447	7.226374	1.330321	0.682297
Including													
Variable	1	1	1	1	1	1	1	1	1	1.225183	5.730145	1.097432	0.387846
Non-increasing	1	1	1	1	1	1	1	1	1.043285	1.508652	5.998332	1.154154	0.466451
Constant	1	1	1	1	1	1	1	1	1.161712	1.546406	5.998332	1.165338	0.470261

Table 2.7: Spearman tests of correlation of ordinal soil-quality measure including soil carbon and soil clay under constant, non-increasing, and variable returns to scale, between median and maximal reference vectors of inputs and outputs (first half) and for different reference levels, among different returns to scale assumptions (second half)

Returns to scale		Median to Maximum
Variable		0.8164 (n=360)
Non-increasing		0.9081 (n=431)
Constant		0.8345 (n=436)
Returns to scale	**Median**	**Maximum**
Constant vs Non-increasing	0.9763 (n=431)	0.8523 (n=452)
Constant vs Variable	0.7927 (n=360)	0.8549 (n=452)
Non-increasing vs Variable	0.7910 (n=360)	0.9767 (n=452)

Note: Spearman tests are all significant at 1% level.

Table 2.8: Tests for significance of soil quality adjustment: Kolmogorov-Smirnov two-samples tests and bootstrapped Li two-samples tests

Returns to scale	p-value	corrected p-value
Kolmogorov-Smirnov tests		
Efficient units included		
Variable	0.494	0.461
Non-increasing	0.997	0.997
Constant	0.547	0.514
Efficient units not included		
Variable	0.037	0.028
Non-increasing	0.800	0.762
Constant	0.096	0.077
Bootstrap Li tests		
Smoothed distribution		
Variable	0.294	
Non-increasing	0.996	
Constant	0.588	
Efficient units not included		
Variable	0.0505	
Non-increasing	0.7985	
Constant	0.111	

Note: The Kolmogorov-Smirnov tests are performed by including the efficient units or truncating the efficient units and then applying the STATA test 'ksmirnov'. The bootstrap Li tests are performed by first, smoothing the distributions in MATLAB or truncating the efficient units, and then applying the routine 'npdeneqtest' implementing Li et al. (2009) in the R 'NP' package by Hayfield and Racine version 0.50-1 with 2000 bootstrap replications. Equality is rejected if p-value is smaller than the significance level desired. If β is considered to be 0.95, in the smoothed distribution tests, resulting p-values do not change qualitatively, and are 0.931 under variable returns, 0.995 under non-increasing returns, and 0.863 under constant returns to scale.

Table 2.9: Tests for separability: two-tailed test statistic (TS) of separable versus non-separable technology at the highest percentile of reference vectors for different returns to scale

		Maximal Reference Level					
No. of Test	TS	0.5%	2.5%	5%	95%	97.5%	99.5%
Variable							
TS1	1.0954	0.9864	1.0011	1.0095	1.0976	1.1088	1.1255
TS2	1.105	0.9523	0.9617	0.9684	1.0441	1.0528	1.0652
TS3	1	1.0062	1.0165	1.0217	1.0729	1.0775	1.0889
TS4	1	0.9914	1.0039	1.011	1.0876	1.0945	1.1066
TS5	1.0530	0.9911	1.0042	1.0105	1.0842	1.0914	1.107
TS6	1.0852	0.9534	0.9688	0.974	1.0397	1.0474	1.0601
Non-increasing							
TS1	1.0915	0.997	1.0136	1.0229	1.1297	1.1414	1.1596
TS2	1.1009	0.9421	0.9563	0.9632	1.0587	1.0679	1.0817
TS3	1	1.0016	1.0142	1.021	1.0895	1.0967	1.1088
TS4	1	0.9917	1.0083	1.0155	1.1116	1.1215	1.1396
TS5	1.0542	0.9902	1.0055	1.0137	1.1066	1.1156	1.1338
TS6	1.0869	0.9475	0.9617	0.9703	1.0554	1.0629	1.0772
Constant							
TS1	1.0861	1.0024	1.0164	1.0251	1.1345	1.1447	1.1622
TS2	1.0973	0.9485	0.96	0.9668	1.0659	1.0758	1.0901
TS3	1	1.0048	1.0158	1.0245	1.0982	1.1045	1.1183
TS4	1	0.9893	1.0071	1.0161	1.1134	1.1228	1.1403
TS5	1.0502	0.9938	1.0074	1.0171	1.1112	1.1219	1.1417
TS6	1.0821	0.9507	0.9644	0.9725	1.0608	1.0698	1.083

Note: The test rejects at different percent significance levels the null hypothesis of separability if the test statistic (TS) in the second column is less than any of the cut-offs below 5% or higher than any of the cut-offs above 95%.

Chapter 3: Bias-corrected nonparametric test for structural separability

3.1 Introduction

Imposing structural separability on the production technology allows aggregation and, in this manner, simplification of economic and econometric analyses. However, incorrectly maintaining separability—among inputs or outputs—causes erroneous measurement of "aggregate" quantities that are not based on physical reality. Doubtful or even erroneous conclusions could result if these "aggregates" were used. A test for structural separability is a necessary step to validate these conclusions.

Methods to test technological structural separability have long been sought in economics (Woodland, 1978; Diewert and Wales, 1995; Banker, 1996; Emran and Alam, 1999; Berndt and Christensen, 1973; Blackorby et al., 1977). Recently, the development of subsampling bootstrapping techniques to test for other structural restrictions, such as returns to scale in Simar and Wilson (2010) and additive aggregation in Schubert and Simar (2010), have contributed to our toolkit for developing empirical tests for the presence of separability. Even more importantly, these methods have been applied to test separability of environmental variables from the production process (Daraio et al., 2010) as required by the data generating process in Simar and Wilson (2007). However, no author has yet established a two-tailed nonparametric separability test that corrects for the bias deriving from the slow convergence of nonparametric efficiency estimators.

In a recent study, Kneip et al. (2015a) propose methods to account for the bias and variance of nonparametric efficiency estimators in approximating the average efficiency of a population of decision making units. Another recent paper (Kneip et al., 2015b) develops tests for hypotheses on the technology: specifically they test for equality of average efficiencies of two different groups of decision makers, they test for returns to scale, and for convexity of the technology.

The present study adapts such methods to develop a test for separability that can be applied to production frontiers approximated by activity-analysis models. This test consists of comparing different means of efficiency estimators calculated on the same sample, either assuming a group of production factors are separable or assuming the production factors are not separable. We use Monte Carlo experiments to examine the statistical properties of the test proposed.

In the empirical illustration we test structural separability of soil characteristics from other production factors using an agricultural data set from Kenyan household farmers. Soil characteristics can be reliably aggregated into a soil-quality measure if the technology is separable. The same procedure can be applied to any economic problem in which one is willing to assume structural separability.

The next section explains the methodological approach by introducing the nonparametric test for structural separability. The third section contains the Monte Carlo evidence and the fourth section demonstrates the method empirically. The fifth section concludes.

3.2 Methodology

This study proposes a methodology to test whether a subvector of inputs is separable from the other inputs and outputs in a production technology. The test conceptually can be represented as follows:[1]

$$\text{Test} \begin{cases} H_0 : & \text{a subvector of inputs is separable;} \\ \\ \text{versus } H_1 : & \text{a subvector of inputs is not separable.} \end{cases} \tag{3.1}$$

Let $\mathbf{x} \in \mathbb{R}_+^P$ denote a vector of inputs, and $\mathbf{y} \in \mathbb{R}_+^Q$ denote a vector of outputs. Let $\mathcal{X}_n = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^n$ denote the sample of n observations available for estimation. Define the production technology $T \subset \mathbb{R}_+^P \times \mathbb{R}_+^Q$ as all input-output combinations technologically feasible:

$$T = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{P+Q} : \mathbf{x}\, \textit{can be used to produce}\, \mathbf{y}\}. \tag{3.2}$$

T satisfies:

A.1: Non emptiness and closeness;

A.2: Weak disposability of outputs, that is, $(\mathbf{x}, \mathbf{y}) \in T \Longrightarrow (\mathbf{x}, \lambda \mathbf{y}) \in T, 0 < \lambda < 1$.

A Farrell output efficiency score characterizes the technology:

$$E(\mathbf{x}, \mathbf{y}) = \max\left\{e \in \mathbb{R}_+ \colon (\mathbf{x}, e\mathbf{y}) \in T\right\}, \tag{3.3}$$

[1] We use inputs to exemplify the methodology but outputs can be similarly aggregated.

if $\exists\, e$ s.t. $(\mathbf{x}, e\mathbf{y}) \in T$ and 0 otherwise, and where $E : \mathbb{R}_+^P \times \mathbb{R}_+^Q \to \mathbb{R}_+$.

By **A.2**

$$E(\mathbf{x}, \mathbf{y}) \geq 1 \Leftrightarrow \mathbf{y} \in T, \tag{3.4}$$

so that $E(\mathbf{x}, \mathbf{y})$ is a complete function representation of the technology.

A subset of inputs $\mathbf{x}^F \in \mathbb{R}_+^F$ are assumed separable from the other production inputs $\mathbf{x}^C \in \mathbb{R}_+^C$ and outputs $\mathbf{y}$, where $F + C = P$. This suggests the possibility of aggregating the separable inputs $\mathbf{x}^F$ into a single measure through the use of an aggregator function $g : \mathbb{R}_+^F \to \mathbb{R}_+$. If the separability assumption is correct, then:

$$E(\mathbf{x}^C, \mathbf{x}^F, \mathbf{y}) = S(\mathbf{x}^C, g(\mathbf{x}^F), \mathbf{y}), \tag{3.5}$$

where $S : \mathbb{R}_+^C \times \mathbb{R}_+ \times \mathbb{R}_+^Q \to \mathbb{R}_+$ is increasing in g, but g does not need to be monotonic in $\mathbf{x}^F$.

The left- and right-hand sides of (3.5) are the non-separable and the separable versions of the technology, respectively. The left-hand side of (3.5) is characterized by (3.3) in one step. On the contrary, the right-hand side of (3.5) is obtained in two steps. From the Farrell efficiency score, in a first step, one can construct directly a monotonic transformation $m(\cdot)$ of the input aggregator function $g(\mathbf{x}^F)$. A fundamental result in separability theory (Theorem 3.2a and Corollary 3.2.0a of Blackorby et al. 1978) is used to define the monotonic transformation of $g(\mathbf{x}^F)$ as follows:

$$m(g(\mathbf{x}^F)) := E(\bar{\mathbf{x}}^C, \mathbf{x}^F, \bar{\mathbf{y}}), \tag{3.6}$$

where $\bar{\mathbf{x}}^C, \bar{\mathbf{y}}$ denote reference levels. The resulting monotonic transformation $m(g(\mathbf{x}^F))$, which is measured in the same units as the efficiency score, can be taken as a summary measure of the aggregated inputs $\mathbf{x}^F$. This measure is similar to the structure in Zieschang (1983) of a scaling function, which results from a Sincov functional equation in logarithmic form, according to Áczel (1970). The second step of the characterization of the separable version of the technology consists in plugging (3.6) into (3.5) to obtain:

$$S(\mathbf{x}^C, m(g(\mathbf{x}^F)), \mathbf{y}) = \max\left\{e^s \in \mathbb{R}_+ \colon (\mathbf{x}^C, m(g(\mathbf{x}^F)), e^s\mathbf{y}) \in T\right\}. \tag{3.7}$$

We divide the remaining of this section into two subsections. In the first subsection we propose an empirical model to identify the left- and right-hand side terms in equality (3.5) to test the separability hypothesis. In the second subsection we introduce a statistical methodology to recognize separable and non-separable technologies. After an introduction of the statistical test, the second subsection unfolds in two parts: in the first, we introduce mathematical notation needed to develop the results. In the second part, we derive the bias-corrected test of difference of means for non independent samples when estimators' convergence rates are different.

3.2.1 An empirical Data Envelopment Analysis model

The technology set T is approximated in this study with a DEA nonparametric piece-wise linear model. For this reason, in addition to the previous assumptions, T satisfies:

A.3: Convexity: If $(\mathbf{x}_1, \mathbf{y}_1) \in T$ and $(\mathbf{x}_2, \mathbf{y}_2) \in T$, then $\forall a \in [0, 1] : a(\mathbf{x}_1, \mathbf{y}_1) + (1 - a)(\mathbf{x}_2, \mathbf{y}_2) \in T$.

A.4: Boundedness of output set: $Y(\mathbf{x}) = \{\mathbf{y} : (\mathbf{x}, \mathbf{y}) \in T\}$, $\forall \mathbf{x} \in \mathbb{R}^P_+$.

A.5: Strong disposability of outputs: If $(\mathbf{x}, \mathbf{y}) \in T$ then $0 \leq \mathbf{y}' \leq \mathbf{y} \Rightarrow (\mathbf{x}, \mathbf{y}') \in T$.

A.6: Strong disposability of inputs: If $(\mathbf{x}, \mathbf{y}) \in T$ then $\mathbf{x}' \geq \mathbf{x} \Rightarrow (\mathbf{x}', \mathbf{y}) \in T$.

Define the approximation to the technology set T as follows:

$$\begin{aligned} \hat{T} = \{y_q &\leq \sum_{i=1}^{n} \zeta_i y_{iq},\ q = 1, \ldots, Q; \\ x_c &\geq \sum_{i=1}^{n} \zeta_i x_{ic},\ c = 1, \ldots, C; \\ x_f &\geq \sum_{i=1}^{n} \zeta_i x_{if},\ f = 1, \ldots, F; \\ \text{for}(\zeta_1, \ldots, \zeta_n)\ \text{s.t.} &\sum_{i=1}^{n} \zeta_i = 1, \zeta_i \geq 0, i = 1, \ldots, n\}, \end{aligned} \tag{3.8}$$

where i indexes each of the n decision making units.[2] The characterization of the non-separable version of the technology (3.3) is empirically estimated through the following DEA linear program:

$$\begin{aligned} \hat{E}(\mathbf{x}, \mathbf{y}) &= \max e \in \mathbb{R}_+ \\ \text{s.t.}\ e y_q &\leq \sum_{i=1}^{n} \mu_i y_{iq},\ q = 1, \ldots, Q; \\ x_c &\geq \sum_{i=1}^{n} \mu_i x_{ic},\ c = 1, \ldots, C; \\ x_f &\geq \sum_{i=1}^{n} \mu_i x_{if},\ f = 1, \ldots, F; \\ \text{for}(\mu_1, \ldots, \mu_n)\ \text{s.t.} &\sum_{i=1}^{n} \mu_i = 1, \mu_i \geq 0, i = 1, \ldots, n. \end{aligned} \tag{3.9}$$

[2] Assumptions on returns to scale (here considered as variable) and on disposability of inputs and outputs can be varied in this approximation and in subsequent linear programs. For example, the separable inputs in the empirical application are assumed non-disposable, so that only convex combinations are included: $x_f = \sum_{i=1}^{n} \zeta_i x_{if},\ f = 1, \ldots, F$.

The separable version of the technology is estimated in two steps. The first step estimates the monotonic transformation $m(g(\mathbf{x}^F))$ defined in (3.6):

$$
\begin{aligned}
\hat{m}(g(\mathbf{x}^F)) &= \max a \in \mathbb{R}_+ \\
\text{s.t. } a\bar{y}_q &\leq \sum_{i=1}^{n} \lambda_i y_{iq},\ q = 1, \ldots, Q; \\
\bar{x}_c &\geq \sum_{i=1}^{n} \lambda_i x_{ic},\ c = 1, \ldots, C; \\
x_f &\geq \sum_{i=1}^{n} \lambda_i x_{if},\ f = 1, \ldots, F; \\
\text{for}(\lambda_1, \ldots, \lambda_n) \text{ s.t. } &\sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0, i = 1, \ldots, n,
\end{aligned} \tag{3.10}
$$

where $\bar{x}_c, \bar{y}_q$ denote reference levels. The second step uses the result of (3.10) to obtain an estimate of (3.7) as follows:

$$
\begin{aligned}
\hat{S}(\mathbf{x}^C, \hat{m}(g(\mathbf{x}^F)), \mathbf{y}) &= \max e^s \in \mathbb{R}_+ \\
\text{s.t. } e^s y_q &\leq \sum_{i=1}^{n} \gamma_i y_{iq},\ q = 1, \ldots, Q; \\
x_c &\geq \sum_{i=1}^{n} \gamma_i x_{ic},\ c = 1, \ldots, C; \\
\hat{m}(g(\mathbf{x}^F)) &\geq \sum_{i=1}^{n} \gamma_i \hat{m}(g(\mathbf{x}_i^F)); \\
\text{for}(\gamma_1, \ldots, \gamma_n) \text{ s.t. } &\sum_{i=1}^{n} \gamma_i = 1, \gamma_i \geq 0, i = 1, \ldots, n.
\end{aligned} \tag{3.11}
$$

3.2.2 The structural separability test

We propose to formalize mathematically our test (3.1) as follows:

$$\text{Test}\begin{cases} H_0: & S(\mathbf{x}^C, m(g(\mathbf{x}^F)), \mathbf{y}) - E(\mathbf{x}^C, \mathbf{x}^F, \mathbf{y}) = 0; \\ \\ \text{versus } H_1: & S(\mathbf{x}^C, m(g(\mathbf{x}^F)), \mathbf{y}) - E(\mathbf{x}^C, \mathbf{x}^F, \mathbf{y}) \neq 0. \end{cases} \tag{3.12}$$

Because we test for separability in the whole technology, we consider relevant the population average measures of S and E: η_S and μ_E respectively. The bias in the approximation to the population mean is of larger magnitude than the variance or the covariance among observations (Kneip et al., 2015a). For this reason, Kneip et al. (2015a) consider the bias explicitly when making inference about the population mean. In this paper we adapt the results about new central limit theorems in Kneip et al. (2015b,a) to the case of separability. Testing separability involves comparing, within the same sample, population means of estimators with different convergence rates.

3.2.2.1 Mathematical notation

We first recognize that the DEA estimates obtained from equations (3.9), (3.10), and (3.11) are empirical estimates of their true counterparts in equations (3.3), (3.6), and (3.7). The same recognition is true also for the estimates of their averages. In addition, the true DEA scores in equations (3.3), (3.6), and (3.7) are only evaluated at random points so we have to evaluate the covariances of the estimators. As the bias is of larger order than the covariance, standard central limit theorems do not hold (Kneip et al., 2015a).

We denote the sample average of the true efficiency measures under non-separability as

$$\bar{E}_n = n^{-1} \sum_{i=1}^{n} E(\mathbf{x}_i, \mathbf{y}_i) \tag{3.13}$$

and let the sample average of the true efficiency measures under separability be

$$\bar{S}_n = n^{-1} \sum_{i=1}^{n} S(\mathbf{x}_i^C, m(g(\mathbf{x}_i^F)), \mathbf{y}_i). \tag{3.14}$$

Moreover, we denote the sample average of the estimated efficiency measures under non-separability as

$$\hat{\mu}_n = n^{-1} \sum_{i=1}^{n} \hat{E}(\mathbf{x}_i, \mathbf{y}_i) \tag{3.15}$$

and let the sample average of the estimated efficiency measures under separability be

$$\hat{\eta}_n = n^{-1} \sum_{i=1}^{n} \hat{S}(\mathbf{x}_i^C, \hat{m}(g(\mathbf{x}_i^F)), \mathbf{y}_i). \tag{3.16}$$

Nonparametric estimators used in this study have different convergence rates depending on the number of inputs and outputs. The convergence rate for the estimator under non-separability depends on n^{κ} where κ varies with the assumptions on the shape of the technology: for example, $\kappa = 2/(C+F+Q+1)$ for the variable returns to scale (VRS) estimator and $\kappa = 2/(C+F+Q)$ for the constant returns to scale (CRS) estimator. On the other hand, the estimator under separability uses lower number of inputs than the corresponding estimator under non-separability. The convergence rate for the estimators under separability depends on n^{κ_α}. Also in this case κ_α varies with the estimator used: for example, $\kappa_\alpha = 2/(C+1+Q+1)$ for the VRS estimator, and $\kappa_\alpha = 2/(C+1+Q)$ for the CRS estimator. Because it uses lower number of inputs, the estimator under separability converges faster than the estimator under non-separability.

3.2.2.2 Bias-corrected test for separability in dependent samples

The Lindeberg-Feller central limit theorem (Kneip et al., 2015a) establishes that each sample estimator has a normal limiting distribution:

$$\sqrt{n}(\bar{E}_n - \mu_E) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_E^2) \tag{3.17}$$

and

$$\sqrt{n}(\bar{S}_n - \eta_S) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_S^2). \tag{3.18}$$

The separability test proposed in this study consists of a difference of means. The difference between sample averages and population means under non-separability and under separability is:

$$(\hat{\eta}_n - \hat{\mu}_n) - (\eta_S - \mu_E). \tag{3.19}$$

This composite difference is in turn equal to the sum of two differences: the difference between averages of the true efficiency scores and the population means, and the difference between averages of estimated and true efficiency measures:

$$\begin{aligned}&(\hat{\eta}_n - \hat{\mu}_n) - (\eta_S - \mu_E) = (\bar{S}_n - \bar{E}_n) - (\eta_S - \mu_E) + \\ &n^{-1}\sum_{i=1}^{n}[(\hat{S}(\mathbf{x}_i^C, \hat{m}(g(\mathbf{x}_i^F)), \mathbf{y}_i) - \hat{E}(\mathbf{x}_i, \mathbf{y}_i)) - (S(\mathbf{x}_i^C, m(g(\mathbf{x}_i^F)), \mathbf{y}_i) - E(\mathbf{x}_i, \mathbf{y}_i))].\end{aligned} \tag{3.20}$$

Under the null hypothesis, the asymptotic variance of this statistic is zero (Kneip et al., 2015b). Such a test statistic has a degenerate distribution under the null: as shown in Kneip et al. (2015b), the test statistic collapses to a Dirac delta function at zero.

We adapt the results in section 3.3 of Kneip et al. (2015b) to obtain non degenerate statistics under the null hypothesis. We split the sample $\mathcal{X}_n$ into mutually exclusive portions $\mathcal{X}_{n_1}$ and $\mathcal{X}_{n_2}$. The non-separable and the separable versions of the estimator are applied respectively to the separate subsamples $\mathcal{X}_{n_1}$ and $\mathcal{X}_{n_2}$. Similarly to section 3.3 in Kneip et al. (2015b), the estimators have different convergence rates under the null: n^{κ} and n^{κ_α}. These rates vary depending on the technology assumption adopted.

Given a certain assumption on the shape of the technology, the estimator calculated under separability converges faster than the estimator under non-separability. To compensate the slower convergence rate, we adapt the intuition in Kneip et al. (2015b), and allocate more observations to the slower converging estimator when splitting the sample. In other words, we set $n_1^{\kappa} = n_2^{\kappa_\alpha}$ and $n_1 + n_2 = n$. With simple numerical methods we can find a solution to the following: $n - n_2 - n_2^{\kappa_\alpha/\kappa} = 0$ and then set n_2 equal to the integer part of the solution and $n_1 = n - n_2$. In particular, $n_1 > n_2$ so that the higher number of observations goes to the non-separable estimator, which has slower convergence rate (Kneip et al., 2015b).

We can then compute the means under non-separability

$$\hat{\mu}_{n_1} = (n_1)^{-1} \sum_{i=1}^{n_1} \hat{E}(\mathbf{x}_i, \mathbf{y}_i \mid \mathcal{X}_{n_1}) \tag{3.21}$$

and under separability

$$\hat{\eta}_{n_2} = (n_2)^{-1} \sum_{i=1}^{n_2} \hat{S}(\mathbf{x}_i^C, \hat{m}(g(\mathbf{x}_i^F)), \mathbf{y}_i \mid \mathcal{X}_{n_2}). \tag{3.22}$$

The sample variances can instead be computed under non-separability as

$$\hat{\sigma}^2_{\mu,n_1} = (n_1)^{-1} \sum_{i=1}^{n_1} [\hat{E}(\mathbf{x}_i, \mathbf{y}_i \mid \mathcal{X}_{n_1}) - \hat{\mu}_{n_1}]^2 \tag{3.23}$$

and under separability as

$$\hat{\sigma}^2_{\eta,n_2} = (n_2)^{-1} \sum_{i=1}^{n_2} [\hat{S}(\mathbf{x}_i^C, \hat{m}(g(\mathbf{x}_i^F)), \mathbf{y}_i \mid \mathcal{X}_{n_2}) - \hat{\eta}_{n_2}]^2. \tag{3.24}$$

The estimators of the mean and variances on split samples, under theorem 4.1 of Kneip et al. (2015a), attain asymptotically the mean and variances of the population. To construct bias corrections for the population means we use K different possible splits of the subsamples of numerosity n_1 and n_2. Number K should be much less than the maximal number of different combinations of split subsamples.[3] For each k^{th} possible split we subdivide the subsamples into two halves indexed by $j \in \{1, 2\}$: m_{1j} and m_{2j}.[4] For the two different subsamples $\mathcal{X}_{n_1}$ and $\mathcal{X}_{n_2}$ respectively we compute, for $j \in \{1, 2\}$:

$$\hat{\mu}^j_{m_{1j},k} = (m_{1j})^{-1} \sum_{i=1}^{m_{1j}} \hat{E}(\mathbf{x}_i, \mathbf{y}_i \mid \mathcal{X}^j_{m_{1j},k}) \tag{3.25}$$

[3]We follow the advice of Kneip et al. (2015b) and fix the number of splits as the minimum between the possible number of combinations and 100. In the subsampling procedure some drawn subsamples provoked infeasibilities in the calculation of the monotonic aggregate measure. These values were replaced by the average measure of the feasible units for calculating the right-hand side of (3.5). This is one way to minimize the bias deriving from this infeasibility problem but one could think of types of more complicated balanced subsampling that potentially allow better bias correction results. However, these methods impose restrictions that are not present in the subsamples. Moreover, as we will show in the analysis of the Monte Carlo experiments, the asymptotic results in this study seem not extremely affected by this problem.

[4]In case of an odd number of units the second half of the subsample contains one more unit.

and

$$\hat{\eta}^{j}_{m_{2j},k} = (m_{2j})^{-1} \sum_{i=1}^{m_{2j}} \hat{S}(\mathbf{x}_i^C, \hat{m}(g(\mathbf{x}_i^F)), \mathbf{y}_i \mid \mathcal{X}^{j}_{m_{2j},k}). \tag{3.26}$$

Then we obtain

$$\hat{\mu}^{*}_{n_1,k} = 0.5 * \{\hat{\mu}^{1}_{m_{11},k} + \hat{\mu}^{2}_{m_{12},k}\} \tag{3.27}$$

and

$$\hat{\eta}^{*}_{n_2,k} = 0.5 * \{\hat{\eta}^{1}_{m_{21},k} + \hat{\eta}^{2}_{m_{22},k}\}. \tag{3.28}$$

For each of the k splits we obtain the generalized jackknife bias estimate under the two different assumptions:

$$\tilde{B}_{\kappa\mu_{n_1},k} = (2^{\kappa} - 1)^{-1}(\hat{\mu}^{*}_{n_1,k} - \hat{\mu}_{n_1}) \tag{3.29}$$

and

$$\tilde{B}_{\kappa_\alpha\eta_{n_2},k} = (2^{\kappa_\alpha} - 1)^{-1}(\hat{\eta}^{*}_{n_2,k} - \hat{\eta}_{n_2}). \tag{3.30}$$

Finally, the averages of the biases over the splits are the bias corrections we consider in the separability test:

$$\hat{B}_{\kappa\mu_{n_1}} = (K)^{-1} \sum_{k=1}^{K} \tilde{B}_{\kappa\mu_{n_1},k} \tag{3.31}$$

and

$$\hat{B}_{\kappa_\alpha\eta_{n_2}} = (K)^{-1} \sum_{k=1}^{K} \tilde{B}_{\kappa_\alpha\eta_{n_2},k}. \tag{3.32}$$

Under the null hypothesis of separability of the technology, and adapting from Kneip et al. (2015b), theorem 4.2 in Kneip et al. (2015a) ensures that:

$$\hat{\tau}_{1,n} = \frac{(\hat{\eta}_{n_2} - \hat{\mu}_{n_1}) - (\hat{B}_{\kappa_\alpha\eta_{n_2}} - \hat{B}_{\kappa\mu_{n_1}})}{\sqrt{\frac{\hat{\sigma}^2_{\mu,n_1}}{n_1} + \frac{\hat{\sigma}^2_{\eta,n_2}}{n_2}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{3.33}$$

The null hypothesis of separability is rejected at an α significance level if $\hat{\tau}_{1,n} < -\tau_{(\alpha/2)}$ or $\hat{\tau}_{1,n} > \tau_{(\alpha/2)}$, where $\tau_{(\cdot)}$ are the corresponding (two-tailed) critical values from the standard normal distribution. This theorem is applicable only if the total number of inputs and outputs is such that both $\kappa_\alpha \geq 2/5$ and $\kappa \geq 2/5$. For example, this restriction means that $C+F+S \leq 4$ under variable returns to scale. However, no meaningful example with multiple inputs ($C > 1$) or outputs ($S > 1$) and more than one aggregated input ($F > 1$) can be conceived in this case of variable returns to scale for such a small dimensionality.

Theorem 4.4 in Kneip et al. (2015a) is applicable if instead the number of inputs and outputs exceed these conditions. Similar to Kneip et al. (2015b), we need to resort to computing the sample means ($\hat{\eta}_{n_2} - \hat{\mu}_{n_1}$) using subsets of the available observations. We subsample two subsets of numerosity $n_{1,\kappa} = [n_1^{2\kappa}]$ and $n_{2,\kappa_\alpha} = [n_2^{2\kappa_\alpha}]$ constituting subsamples $\mathcal{X}^*_{1,n_{1,\kappa}}$ and $\mathcal{X}^*_{2,n_{2,\kappa_\alpha}}$ composed of random subsets of elements from the subsamples $\mathcal{X}_{n_j}$ with $j \in \{1,2\}$.[5] From these subsamples we obtain

$$\hat{\mu}_{n_{1,\kappa}} = (n_{1,\kappa})^{-1} \sum_{(\mathbf{x}_i,\mathbf{y}_i)\in\mathcal{X}^*_{1,n_{1,\kappa}}} \hat{E}(\mathbf{x}_i, \mathbf{y}_i \mid \mathcal{X}_{n_1}) \tag{3.34}$$

under non-separability and

$$\hat{\eta}_{n_{2,\kappa_\alpha}} = (n_{2,\kappa_\alpha})^{-1} \sum_{(\mathbf{x}_i^C, \hat{m}(g(\mathbf{x}_i^F)), \mathbf{y}_i)\in\mathcal{X}^*_{2,n_{2,\kappa_\alpha}}} \hat{S}(\mathbf{x}_i^C, \hat{m}(g(\mathbf{x}_i^F)), \mathbf{y}_i \mid \mathcal{X}_{n_2}) \tag{3.35}$$

under separability assumption. Under the null hypothesis of separability of the

[5]We let the symbol $[a]$ denote the integer part of a. Monte Carlo experiments and empirical analyses have been repeated setting κ_α equal to κ and results are qualitatively similar.

technology, theorem 4.4 of Kneip et al. (2015a) ensures the following:

$$\hat{\tau}_{2,n} = \frac{(\hat{\eta}_{n_2,\kappa_\alpha} - \hat{\mu}_{n_1,\kappa}) - (\hat{B}_{\kappa_\alpha \eta_{n_2}} - \hat{B}_{\kappa \mu_{n_1}})}{\sqrt{\frac{\hat{\sigma}^2_{\mu,n_1}}{n_{1,\kappa}} + \frac{\hat{\sigma}^2_{\eta,n_2}}{n_{2,\kappa_\alpha}}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{3.36}$$

In a similar manner to $\hat{\tau}_{1,n}$, we reject the null hypothesis of separability at an α significance level if $\hat{\tau}_{2,n} < -\tau_{(\alpha/2)}$ or $\hat{\tau}_{2,n} > \tau_{(\alpha/2)}$, where $\tau_{(\cdot)}$ are the corresponding (two-tailed) critical values from the standard normal distribution.

Depending on the number of inputs and outputs, either $\hat{\tau}_{1,n}$ or $\hat{\tau}_{2,n}$ can be used to test the separability assumption on the technology against standard normal critical values. To ensure the relevance of the test in real-world separability problems, we concentrate on $\hat{\tau}_{2,n}$. In the Monte Carlo experiments and in the empirical illustration we adopt two alternative definitions of the mean difference $(\hat{\eta}_{n_2,\kappa_\alpha} - \hat{\mu}_{n_1,\kappa})$ in order to see how much the results diverge. One alternative is to subsample subsets $\mathcal{X}^*_{1,n_{1,\kappa}}$ and $\mathcal{X}^*_{2,n_{2,\kappa_\alpha}}$ from the subsamples $\mathcal{X}_{n_j}$, without repetition, until all possible subsample combinations are drawn. The averages of those combinations for the separable and non-separable technologies are used as means in the test, instead of the subsample means ($\hat{\eta}_{n_2,\kappa_\alpha}$ and $\hat{\mu}_{n_1,\kappa}$). As a similar alternative, we also use the simple means $\hat{\eta}_{n_2}$ and $\hat{\mu}_{n_1}$ in place of the means from the subsamples $\hat{\eta}_{n_2,\kappa_\alpha}$ and $\hat{\mu}_{n_1,\kappa}$.

3.3 Monte Carlo evidence

The Monte Carlo experiments are useful to analyze the performance of the tests in terms of size and power. According to statistical theory, the size of the test is studied under the null hypothesis, while its power is analyzed for scenarios increasingly diverging from the null hypothesis.

We consider a production technology that transforms three inputs into two outputs:

$$\psi(y_1, y_2) = t(x_1, x_2, x_3), \tag{3.37}$$

where $t : \mathbb{R}^3_+ \to \mathbb{R}^2_+$. To simulate the outputs, we approximate the technology via a transcendental logarithmic (or translog) separability-flexible production function with three inputs:

$$lny_1 = -lny_2 + a_{11}lnx_1 + a_{12}lnx_2 + a_{21}lnx_3 +$$

$$1/2(b_{11}lnx_1lnx_1 + b_{12}lnx_1lnx_2 + b_{13}lnx_1lnx_3 + b_{21}lnx_2lnx_1) +$$

$$1/2(b_{22}lnx_2lnx_2 + b_{23}lnx_2lnx_3 + b_{31}lnx_3lnx_1 + b_{32}lnx_3lnx_2 + b_{33}lnx_3lnx_3). \tag{3.38}$$

The Monte Carlo experiments are used to test the separability of inputs (x_1, x_3) from the other input x_2. The separability hypothesis suggests the following production structure:

$$\psi(y_1, y_2) = t_1(r(x_1, x_3), x_2), \tag{3.39}$$

where $r : \mathbb{R}^2_+ \to \mathbb{R}_+$ is an aggregator function. As shown in the methodology, the separability test can be expressed as:[6]

$$\text{Test} \begin{cases} H_0 : & S_M(x_2, m(r(x_1, x_3)), y_1, y_2) - E_M(x_2, x_1, x_3, y_1, y_2) = 0; \\ \text{versus } H_1 : & S_M(x_2, m(r(x_1, x_3)), y_1, y_2) - E_M(x_2, x_1, x_3, y_1, y_2) \neq 0, \end{cases} \tag{3.40}$$

where S_M and E_M are the Monte Carlo counterparts of S and E, respectively.

[6]The notion of separability employed is a non-symmetric notion of separability.

The three inputs in the Monte Carlo experiments are drawn as uniform random variables:

$$x_1, x_2, x_3 \sim U(1.1, 10). \tag{3.41}$$

The second output is also drawn as a uniform random variable:

$$lny_2 \sim U(1.1, 14). \tag{3.42}$$

By varying the parameters of the translog production function one obtains a separable and a non-separable structure. In particular, for a decreasing returns to inputs production function, the parameters are as follows: $a_{11} = 0.25$, $a_{12} = 0.05$, $a_{21} = 0.25$, $b_{11} = -0.5$, $b_{33} = -b_{13} - b_{23}$, $b_{13} = -b_{12} - b_{11}$, $b_{12} = 1 * \delta$, $b_{23} = 0.0001 * \delta$, and $b_{22} = -b_{12} - b_{23}$. The parameters are chosen as follows for a constant returns to inputs production function: $a_{11} = 0.5$, $a_{12} = 0.05$, $a_{21} = 0.45$, $b_{11} = -0.5$, $b_{33} = -b_{13} - b_{23}$, $b_{13} = -b_{12} - b_{11}$, $b_{12} = 1 * \delta$, $b_{23} = 0.0001 * \delta$,and $b_{22} = -b_{12} - b_{23}$.

If we impose $\delta = 0$, then the technology in (3.39) is separable; for increasing departures from this level in the positive orthant the technology is not separable. These parameters result in a technology homothetic in all inputs. Even if we assume that the technology is homothetic in all inputs, the structure in (3.5) would only be separable if the order of the measures in the monotonic transformation $m(g(\mathbf{x}^F))$ were maintained while changing all non-separable inputs in the same proportion.

We show intuitively the definition of separability in figure 3.1 with the decreasing returns case. We represent two series of isoquants drawn by fixing x_2 at two different levels: solid isoquants for $x_2 = 1.1$ and $y_2 = 1.1$, and dotted isoquants for $x_2 = 10$ and $y_2 = 1.1$. Because the isoquants among the two inputs x_1 and x_3 are parallel, when changing the other input x_2, the two inputs x_1 and x_3 are separable from the third, given the definition in Blackorby et al. (1978).

Varying the parameters of the translog production function obtains a non-separable structure. In particular, increasing δ from 0 to positive numbers increases the non-separability of the production technology.

To show a conservative example of a non-separable technology, in figure 3.2 we fix $\delta = 0.25$. The isoquants among inputs x_1 and x_3 are not parallel, while fixing the other input and output at $x_2 = 1.1$ and $y_2 = 1.1$ (solid isoquants), and at $x_2 = 10$ and $y_2 = 1.1$ (dotted isoquants). Consequently, x_1 and x_3 are not separable from x_2 in this case (Blackorby et al., 1978).

Finally, we simulate an inefficiency process for one of the outputs to obtain a simulated realized output y_{1r}.[7] Calculated output y_1 in equation (3.38) is multiplied by an efficiency term $exp(- \mid u \mid)$:

$$y_{1r} = exp(ln(y_1)) * exp(- \mid u \mid), \tag{3.43}$$

where u is a normal randomly distributed term $u \sim N(0, 0.25)$.

Size and power properties of the tests proposed in this study are explored in different experiments. Each experiment consists of one analysis under the null hypothesis (separable structure), and two under the alternative hypothesis (non-separable structure), for increasing departures from the null. Each analysis consists of calculations for 1000 Monte Carlo trials.[8] On each Monte Carlo trial the tests

[7]We wish to maintain the realized outputs as close as possible to the simulated outputs. For this reason, we consider inefficiency in only one output.

[8]Monte Carlo simulations have been performed on the Cluster Of Unix Machines (CLOU) at

are performed under the same reference level for input x_2, and outputs y_{1r} and y_2. This reference level is fixed in these experiments at the 99^{th} percentile. To show the properties of the test, we consider four sample sizes: 50, 200, 450, and 1000 units.

An increasing power of the test prescribes that rejection rates increase progressively from the separable to the non-separable cases. To study this property, we vary the technology parameters along a specific expansion path by modifying δ.[9] In particular, we vary b_{12} and b_{23} along the following path: $(b_{12} = 1*\delta, b_{23} = 0.0001*\delta)$. We vary δ with two-units increments three times to create three experiments along the same path, corresponding to increasing departures from the null hypothesis of separability: δ_0: $(b_{12} = 0, b_{23} = 0)$, δ_2: $(b_{12} = 2, b_{23} = 0.0002)$, and δ_4: $(b_{12} = 4, b_{23} = 0.0004)$. The separability hypothesis is maintained under scenario δ_0.

The results of the Monte Carlo tests are presented in tables 3.1 and 3.2. In table 3.1 we present the results under decreasing returns to inputs and in table 3.2 we report the results under constant returns to inputs. Each table is divided in three panels and each panel is divided in three sections depending on the scenario considered (δ_0, δ_2, and δ_4). Each panel corresponds to a different way of calculating the mean difference $(\hat{\eta}_{n_2,\kappa_\alpha} - \hat{\mu}_{n_1,\kappa})$ in equation (3.36).

We put in the central panel (panel 2) the results when the mean difference is calculated by sampling only once a subsample combination of size n_{2,κ_α} and $n_{1,\kappa}$, as

Humboldt University in Berlin on eight cores. Calculations for Monte Carlo experiments and the empirical illustration in this article have been done with MATLAB software package from MathWorks.

[9] Professor Dr. Léopold Simar suggested me this idea.

written in equation (3.36). In panel 1 the mean difference is calculated by sampling, without repetition, all possible subsample combinations of size n_{2,κ_α} and $n_{1,\kappa}$. In panel 3 the mean difference is calculated by averaging all observations n_2 and n_1, which is similar to the method used in panel 1. In each of three panels, first section represents the separability scenario (δ_0), while the second (δ_2) and third (δ_4) sections represent the test under increasing departures from the null hypothesis. Each two-tailed test is repeated at three nominal confidence levels: 90%, 95%, and 99%.

In the first section of panel 1 of table 3.1, under separability, we see that the test decreases rapidly rejection rates when increasing sizes from 50 to 200 observations. The test attains approximately nominal sizes for large sample sizes, especially above 450 observations for a nominal level of 99% (0.051 with 450 observations and 0.028 with 1000 observations). A slight over-rejection of the nominal sizes makes the test conservative.

The approximation to nominal levels of rejection is worse, especially at low sample sizes, if we look at the first section of panel 2 of table 3.1. This is a scenario also under separability but where the means are calculated only on one random subsample. The approximation is better if we consider the first section of panel 3 of the same table where the mean differences are calculated by averaging all observations in the samples n_2 and n_1. Rejection rates, in panel 3, are very similar to those in panel 1. It is interesting to notice that the divergence between rejection rates in panel 1 and 3 on one hand, and in panel 2 on the other hand, worsens for increasing sample size under the null hypothesis.

In the second section of panel 1 of table 3.1 we study the first non-separability scenario. Indeed, in this set of simulations the rejection rates are increasing when increasing sample size, particularly for sample sizes larger than 200 observations. The levels of rejection reach up to more than 71% for a sample size of 1000 observations. Even though nominal rejection rates are not reached in this set of simulations, we see that rejection rates increase as we increase sample size. This behavior evidences that the statistical test is consistent. In other words, if we were to increase the sample size indefinitely, the nominal rejection rates could be achieved.

The results in second sections of panel 2 and 3 of table 3.1 are very similar to the ones in panel 1. Compared to the separable case, under the alternative hypothesis the results are more similar across panels. We notice, however, slightly more volatile results in central panel 2 than in the other panels.

In the third section of panel 1 of table 3.1 we increase the departure from the null hypothesis. In this set of simulations rejection rates are also increasing steadily for greater sample sizes. Rejection levels are up to more than 78% in this section. Again the test shows consistency when increasing sample sizes. Similar rejection rates are visible in third section of panels 2 and 3, even though the rejection rates in panel 2 are slightly worse than the other two sets in panels 1 and 3.

The three sections in all panels of table 3.1 evidence that increasing departures from the null hypothesis of separability show increasing rejection rates in the test, across all sample sizes and across all nominal confidence levels. Rejection rates indicate the power of the test increases for increasing departures from the null hypothesis.

The results are similar if we consider table 3.2. In table 3.2 we have the simulations for the constant returns to inputs case. In first section of panel 1 of table 3.2, under the separability hypothesis, rejection rates near to nominal are reached if we increase size from very small samples. However, under separability we notice rejection rates higher than in the decreasing returns to inputs case in first sections of table 3.1. Moreover, first section of panel 2 of table 3.2 shows higher rejection rates than in the corresponding sections in panel 1 and 3 of the same table. This fact indicates once more that calculating the means by one simple subsampling could produce more volatile results.

In second and third sections of table 3.2, for increasing departures from the null hypothesis, the test increasingly rejects the false hypothesis correctly. Similar, but generally lower, rejection rates than in the decreasing returns to inputs case of table 3.1, are evidenced at all sample sizes and at all nominal levels. However, the highest rate of rejection in table 3.2 is, for a 90% nominal size, approximately 79%, which is higher than the corresponding rejection rate in table 3.1.

From the Monte Carlo experiments we notice that the test is generally consistent for increasing sample sizes and it correctly recognizes a separable from a non-separable technology in the majority of the cases for different technologies. The size properties of the test are more precise than its power properties.

These results show that the practitioner shall be cautious in using a specific nominal size when not rejecting the null. If the practitioner wants to be conservative and is usually content with a 95% level of confidence, in this test one might want to reject already the null if the hypothesis is rejected with 90% level of confidence, considering the under-rejection under the alternative hypothesis and the over-rejection under the null hypothesis.

3.4 Empirical illustration

In the empirical illustration of the estimator properties we employ a random sample of household farmers from 61 sublocations in Kenya. Data are from early 2007 and refer to preceding short and long growing seasons. The survey is named "Research on Poverty, Environment and Agricultural Technologies (REPEAT): Panel studies in Africa". Survey data are obtained from the National Graduate Institute for Policy Studies (21st century Center of Excellence Program) in Japan. The present cross-section sample is composed of 590 household farmers.[10] Of these, data on soil quality are available only for 452 families.[11]

Variables used in this analysis are on 17 inputs and 4 outputs. Table 3.3 shows input and output summary statistics. The analysis is done at the household and not at the farm level to account for off-farm income, which is a joint output.

[10] More household farmers are included in the original sampling scheme but are left out of the analysis because have incomplete or erroneous entries.

[11] We checked through a Probit regression and no selectivity bias seems relevant in this sample reduction step for the outputs, which are the metric for measuring the efficiency scores. In the case of the inputs, the resulting sample might suffer some bias: farmers with lower number of cows are more probable to be observed in the final sample. Farmers who spend more in hired labor, and use more organic, and less inorganic, fertilizers are observed more often. These apparent biases are not critical for the methodological results of this paper. They should however be considered when evaluating the representativity of our Kenyan sample. Data have undergone outlier detection methods, following Wilson (1993), to diminish at its minimum measurement error problems.

Main agricultural output considered is harvested dry maize, which is totally rain-fed. All present farmers produce maize. The measured inputs used in maize production are seeds, land area, organic and inorganic fertilizers, family worked hours, cost of temporary hired workers, and hours worked by permanent and shared workers.[12] Other agricultural inputs are number of hand hoes, ploughs, sickles, spray pumps, and household members available to work in the family.[13]

As diversification options other representative outputs are harvested green maize, milk, and off-farm income. Measured inputs available for household production are livestock permanent laborers, number of milking cows, and family members engaged in off-farm activities.

---

[12]Faithful to the human capital approach pioneered by Jorgenson and Griliches (1967), labor input is adjusted for differences in quality, both due to education and age. To adjust labor for education, estimates of the impact of education on agricultural productivity in Kenya from the analysis by Husbands et al. (1996) are used. Given an 85% probability that a primary school-completer household head increases household profitability by 40%, we divide the expected return equally among years of primary education (8 years) to obtain a 4.25% average increase in productivity per additional year of any level of education. It is possible that different methods of partitioning the increase in productivity would have achieved a different result. But this topic is not the focus of the present contribution. The same is done for household farmers with a primary school completer, who is not the household head (increase productivity by 29.75%). To adjust labor for age, hours worked by children who, by United Nations Children Fund classification, are considered those less than 15 years old, are divided in half.

[13]All of the family members living at the household that are above three years old are included. Most kids start completing tasks in the family at this age. Number of members present in the household is also weighted for quality as the number of worked hours.

Data on physical characteristics of soil for the largest maize plot of each household are available for mid-2003. To visualize the method, our analysis focuses mainly on two critical measures of soil structure and texture: soil carbon and soil clay content.[14] Both soil carbon and soil clay content are immutable for long periods unless lands are converted to agriculture immediately before being analyzed.

Soil carbon content (in percentage points of soil weight) is a generally accepted indicator of long-term soil structure (Brady and Weil, 2008) because it varies as a result of human action, but only in the long-run. An increasing number of studies have assumed that more soil carbon content leads to better soil quality because it increases the amount of nutrients available to the plants, and increases infiltration and water-holding capacity. But a too high percentage of soil carbon can interfere with the nitrogen cycle (Brady and Weil, 2008) and be actually indicative of poor drainage of soils. Hence, in our analysis, we allow for the possibility that negative marginal effects may arise from too high levels of carbon content.

Soil clay content (in percentage points of soil weight) provides an important indicator of soil texture. Finer soil clay particles magnify nutrients' retention in the soil (especially nitrogen) for better plant growth. The types of clay on which most household farmers reside (montmorillonitic and kaolinitic) have potentially congestive effects. For both soil carbon and soil clay we adapt hypothesis **A.6** in the linear programs to include only convex combinations of soil characteristics in the technology.

[14] As can be seen in the methodological section, the method is general and would allow for more soil characteristics to be included.

3.4.1 Empirical results

The empirical results obtained from the illustration are reported under variable returns to scale in tables 3.4-3.5, and in figure 3.3. We follow the same convention in presenting the results in the different panels of tables 3.4 and 3.5: in panel 1, means in the test statistic are created by sampling, without replacement, over all possible subsample combinations. In panel 2, means in the test statistic are obtained by sampling only one subsample combination. In panel 3, means are calculated using all observations in each subsample. The statistics and their corresponding probability values (in the adjacent columns) have to be compared to the critical values from the standard normal distribution.

Different statistics are calculated by keeping the reference levels at different percentiles of the empirical distribution of non-separable factors. We perform the test for every percentile of reference levels, from the 49^{th} to the 100^{th} percentile.[15] Some units are not feasible depending on the reference level: the number of infeasibilities, reported in the last column of tables 3.4 and 3.5, decreases monotonically from 64 at the 49^{th} percentile to 0 at the 100^{th} percentile level. The different test results are also graphed in figure 3.3 with corresponding 90% critical values.

The test results show that, on average, the soil characteristics are separable from the other factors at all reference levels considered. All different versions of the two-tailed test statistics are consistent in not rejecting separability, largely within a

[15]By varying the reference levels, we vary the dependence of the monotonic aggregated measure on the reference levels of non-separable inputs and outputs.

90% confidence level.

The probability values signal that the tests never reject significantly the null hypothesis. In the first and third panels the probability values go from 0.215 to 0.414 depending on the reference level. In the second central panel, the probability values are even larger ranging from 0.32 to 0.495. From this analysis we conclude that the test equivalent to equation (3.36) in the central panel would reject the least the separability of the technology.

The test statistics show a parallel behavior in the three panels of tables 3.4-3.5, and in figure 3.3. The test statistics decrease from the 49^{th} percentile until the 56^{th} percentile, then increase up to the 63^{rd} percentile, to fluctuate around those levels until the 70^{th} percentile. It is interesting to notice that test statistics are very stable between 71^{st} percentile and 86^{th} percentile of reference. Finally, test statistics fluctuate more after the 87^{th} percentile but always largely within 90% critical values.[16]

The tests obtained by drawing without replacement all possible subsamples (in panel 1 and depicted with a dashed line in figure 3.3) are very similar (to the third decimal digit) to the tests statistics obtained by averaging all observations in the sample (in panel 3 and depicted by a square diamond in figure 3.3). The expected equality between these two different ways of obtaining the tests and the parallel behavior of these with those in panel 2 (depicted by a dotted line in figure

[16]Instead of dividing the sample in first and second halves, as explained in the theoretical part, we also repeated the test presented by including in each subsample every second observation in order to preserve subsample independence. The results are qualitatively the same.

3.3) show that the test is robust in not rejecting separability.

3.5 Conclusions

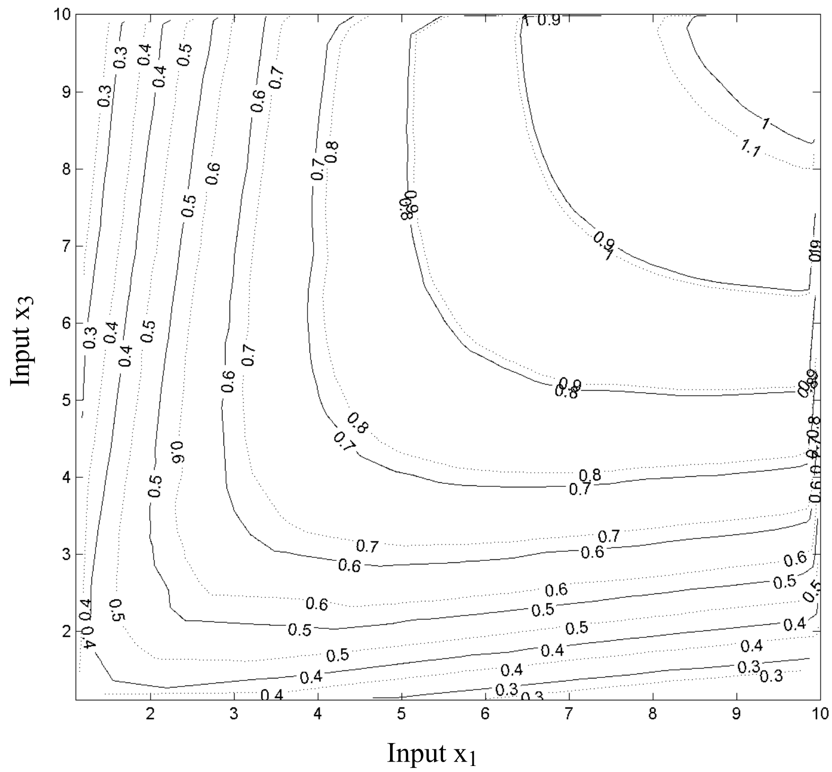
This study proposes a test for structural separability directly among production factors with nonparametric methods. We test the separability of a subvector of production factors (in our case inputs) from the other production factors. Monte Carlo experiments are conducted to show the properties of our test at certain sample sizes. The size properties of the test are more precise than its power properties. However, considering the over-rejection under the null hypothesis and the under-rejection under the alternative hypothesis, the practitioner can impose lower critical values than usual to obtain conservative significance levels.

In our empirical illustration, we apply this testing methodology to a data set of Kenyan household farmers. We test whether soil carbon and soil clay can be assumed separable from other inputs and outputs in a household farmer technology. We test the technological separability under the most general variable returns to scale. The results from our empirical test are clear in not rejecting separability at any meaningful level of significance.

This methodology is applicable in any instance in economics where separability is assumed. For example, one could think of testing, without parametric assumptions, the aggregation of different capital and labor types into capital and labor aggregates. One could also measure the importance of climate change on the production process through the usage of a separable climate indicator over time. The reliability of the separability assumption could be tested with the present methods. In the case of a constant returns to scale aggregator function, the results of the tests proposed could be moreover comparable to nonparametric methodologies of creating separable aggregates under homothetic separability (Lewbel and Linton, 2007).

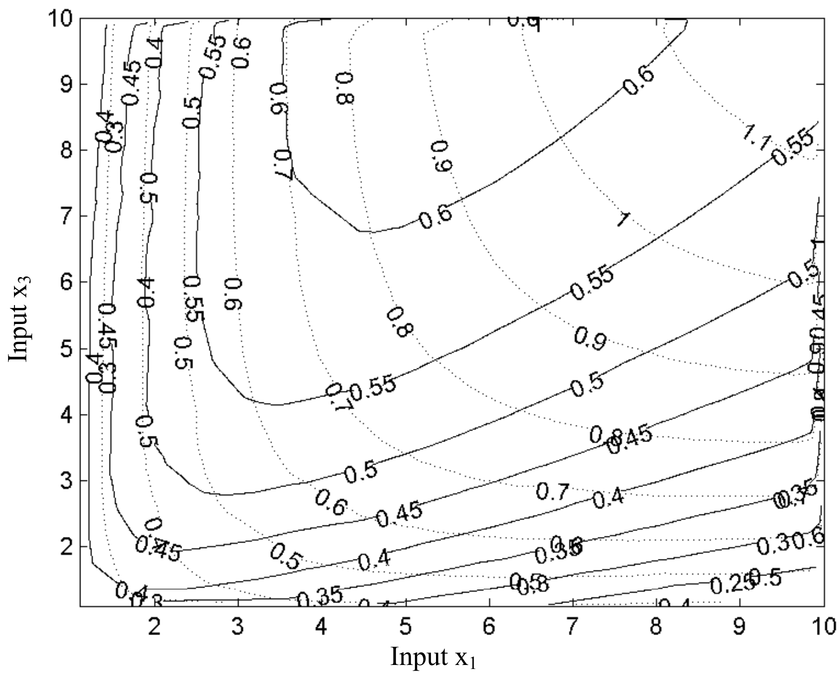
3.6 Figures and Tables

Figure 3.1: Separable translog production function isoquants



Note: The isoquants result from varying inputs x_1 and x_3 reach different values of outputs (depicted in the contour lines), when fixing $x_2 = 1.1$ and $y_2 = 1.1$ in the solid isoquants, and $x_2 = 10$ and $y_2 = 1.1$ in the dotted isoquants.

Figure 3.2: Non-separable translog production function isoquants



Note: The isoquants result from varying inputs x_1 and x_3 and reach different values of outputs (depicted in the contour lines), when fixing $x_2 = 1.1$ and $y_2 = 1.1$ in the solid isoquants, and $x_2 = 10$ and $y_2 = 1.1$ in the dotted isoquants.

Figure 3.3: Separability test on Kenyan household farmers for different percentiles under variable returns to scale

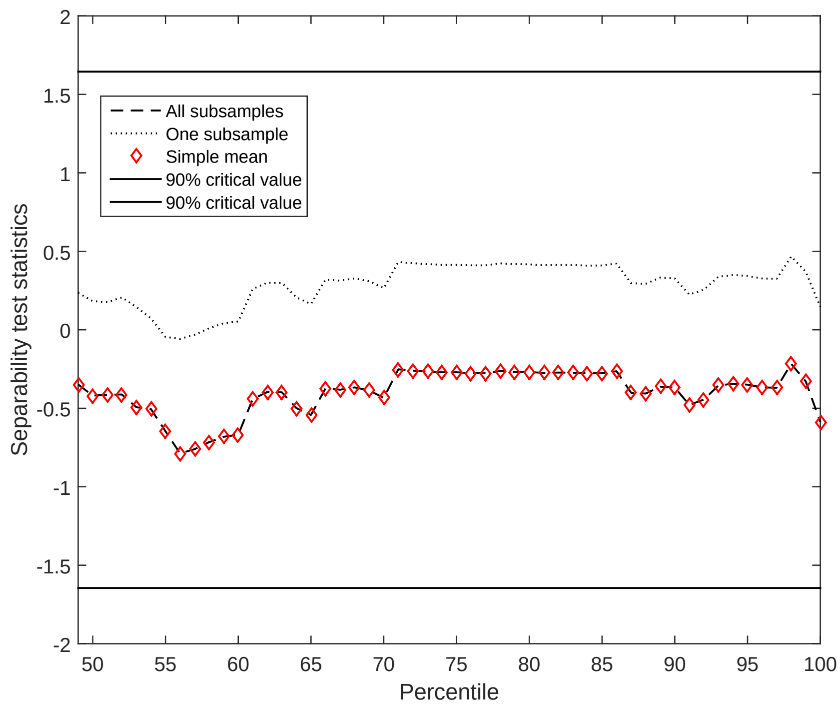


Table 3.1: Rejection rates from Monte Carlo results for a decreasing returns to inputs simulated function

Panel	1: All combinations Nominal levels			2: One sampling Nominal levels			3: Simple mean Nominal levels		
Section I: δ_0									
Observations	99%	95%	90%	99%	95%	90%	99%	95%	90%
50	0.295	0.439	0.526	0.288	0.418	0.5	0.289	0.421	0.51
200	0.067	0.179	0.27	0.129	0.253	0.328	0.073	0.183	0.267
450	0.051	0.12	0.18	0.093	0.19	0.269	0.043	0.118	0.172
1000	0.028	0.094	0.165	0.086	0.188	0.257	0.029	0.099	0.161
Section II: δ_2									
Observations	99%	95%	90%	99%	95%	90%	99%	95%	90%
50	0.396	0.536	0.604	0.397	0.526	0.611	0.382	0.513	0.592
200	0.394	0.526	0.604	0.397	0.534	0.599	0.397	0.528	0.594
450	0.479	0.583	0.646	0.471	0.583	0.653	0.473	0.592	0.649
1000	0.536	0.654	0.716	0.55	0.643	0.704	0.533	0.656	0.717
Section III: δ_4									
Observations	99%	95%	90%	99%	95%	90%	99%	95%	90%
50	0.542	0.651	0.706	0.543	0.644	0.703	0.537	0.656	0.701
200	0.623	0.719	0.764	0.608	0.711	0.771	0.61	0.716	0.768
450	0.656	0.739	0.783	0.647	0.715	0.758	0.658	0.735	0.772
1000	0.646	0.747	0.784	0.632	0.734	0.775	0.645	0.743	0.781

Table 3.2: Rejection rates from Monte Carlo results for a constant returns to inputs simulated function

Panel	1: All combinations Nominal levels			2: One sampling Nominal levels			3: Simple mean Nominal levels		
Section I: δ_0									
Observations	99%	95%	90%	99%	95%	90%	99%	95%	90%
50	0.321	0.459	0.536	0.302	0.43	0.514	0.308	0.443	0.517
200	0.092	0.196	0.28	0.15	0.276	0.351	0.093	0.195	0.277
450	0.068	0.156	0.236	0.114	0.223	0.314	0.07	0.15	0.221
1000	0.079	0.188	0.274	0.138	0.259	0.349	0.083	0.184	0.275
Section II: δ_2									
Observations	99%	95%	90%	99%	95%	90%	99%	95%	90%
50	0.389	0.524	0.608	0.393	0.528	0.615	0.373	0.516	0.584
200	0.349	0.49	0.564	0.366	0.485	0.573	0.358	0.493	0.554
450	0.438	0.552	0.621	0.442	0.543	0.631	0.441	0.551	0.615
1000	0.507	0.622	0.687	0.522	0.63	0.689	0.507	0.619	0.69
Section III: δ_4									
Observations	99%	95%	90%	99%	95%	90%	99%	95%	90%
50	0.529	0.649	0.708	0.531	0.64	0.703	0.527	0.639	0.698
200	0.604	0.711	0.758	0.6	0.711	0.759	0.605	0.713	0.756
450	0.654	0.738	0.783	0.642	0.713	0.761	0.652	0.726	0.767
1000	0.637	0.742	0.791	0.645	0.726	0.772	0.637	0.741	0.791

Table 3.3: Summary statistics of inputs, outputs, and soil quality physical characteristics

Variable	Mean	Std.Dev.	Min	Max
Inputs				
land area (acres)	1.6	1.4	0.1	14
quantity of seeds (kgs)	13.4	11	1	78
inorganic fertilizers (kgs)	46.8	74.4	0	650
organic fertilizers (kgs)	742.1	1214.6	0	9000
hired labor (cost in KSh)	2935.4	4911.9	0	48160
family labor (hours)	431.5	510.9	0	4434.8
permanent and share labor (hours)	41.7	95.1	0	963
number of hand hoes	3.9	2.2	0	15
number of ploughs	0.1	0.3	0	2
number of spray-pumps	0.4	0.6	0	2
number of sickles	0.4	0.6	0	3
livestock permanent labor	0.2	0.7	0	4.8
off-farm earners	1.1	1.1	0	6
present members	4.6	2.7	0	17.5
milking cows	1	0.9	0	5
Outputs				
total harvest dry maize (kgs)	843.7	1122.8	0	9000
total harvest green maize (kgs)	249.4	587.6	0	7180
milk (liters)	1690.9	2435.9	0	18600
off-farm income (in KSh)	85144.4	146018.3	0	1572400
Soil quality physical characteristics				
soil carbon content (% of soil weight)	2.6	1.5	0.7	15.2
soil clay content (% of soil weight)	28.3	3.9	15.5	44.9
Observations	452			

Table 3.4: Separability test empirical results for different reference percentiles from 49^{th} to 75^{th}

Separability Test	H0: technology is separable H1: technology is not separable						
	Panel 1		Panel 2		Panel 3		
Percentile	Statistic	P-value	Statistic	P-value	Statistic	P-value	Infeasible
49	-0.34843	0.36376	0.23559	0.40688	-0.34851	0.36373	64
50	-0.41920	0.33753	0.18277	0.42749	-0.41925	0.33752	63
51	-0.41390	0.33947	0.17638	0.43000	-0.41395	0.33946	61
52	-0.41223	0.34008	0.20650	0.41820	-0.41228	0.34007	61
53	-0.49338	0.31087	0.14565	0.44210	-0.49339	0.31087	55
54	-0.50477	0.30686	0.07357	0.47068	-0.50478	0.30686	55
55	-0.64636	0.25902	-0.04536	0.48191	-0.64632	0.25904	52
56	-0.78763	0.21546	-0.05755	0.47705	-0.78755	0.21548	48
57	-0.76048	0.22348	-0.03188	0.48728	-0.76040	0.22351	47
58	-0.71591	0.23702	0.01014	0.49595	-0.71584	0.23705	46
59	-0.68159	0.24775	0.04209	0.48321	-0.68153	0.24777	45
60	-0.66861	0.25187	0.05312	0.47882	-0.66856	0.25189	42
61	-0.44046	0.32980	0.26039	0.39728	-0.44048	0.32979	41
62	-0.39654	0.34585	0.30072	0.38181	-0.39657	0.34584	38
63	-0.39830	0.34520	0.29893	0.38250	-0.39834	0.34519	37
64	-0.49908	0.30886	0.20605	0.41838	-0.49908	0.30886	26
65	-0.54278	0.29364	0.16441	0.43470	-0.54277	0.29364	25
66	-0.37290	0.35461	0.32072	0.37421	-0.37295	0.35459	21
67	-0.38178	0.35131	0.31304	0.37713	-0.38183	0.35129	21
68	-0.36711	0.35677	0.32752	0.37164	-0.36716	0.35675	21
69	-0.38583	0.34981	0.31041	0.37813	-0.38588	0.34979	20
70	-0.43412	0.33210	0.26530	0.39539	-0.43415	0.33209	20
71	-0.25190	0.40056	0.43186	0.33292	-0.25199	0.40053	17
72	-0.25989	0.39748	0.42444	0.33562	-0.25997	0.39744	17
73	-0.26638	0.39497	0.41843	0.33782	-0.26646	0.39494	17
74	-0.27065	0.39333	0.41439	0.33930	-0.27074	0.39330	17
75	-0.27059	0.39335	0.41456	0.33923	-0.27067	0.39332	17

Table 3.5: Separability test empirical results for different reference percentiles from 76^{th} to 100^{th}

Separability Test	H0: technology is separable						
	H1: technology is not separable						
	Panel 1		Panel 2		Panel 3		
Percentile	Statistic	P-value	Statistic	P-value	Statistic	P-value	Infeasible
76	-0.27640	0.39112	0.41090	0.34057	-0.27648	0.39109	14
77	-0.27678	0.39098	0.41050	0.34072	-0.27686	0.39095	14
78	-0.26259	0.39643	0.42341	0.33600	-0.26268	0.39640	13
79	-0.26791	0.39438	0.41846	0.33781	-0.26799	0.39435	13
80	-0.26936	0.39383	0.41690	0.33838	-0.26945	0.39379	13
81	-0.27448	0.39186	0.41165	0.34030	-0.27457	0.39183	13
82	-0.27251	0.39262	0.41360	0.33959	-0.27259	0.39258	13
83	-0.27311	0.39238	0.41308	0.33977	-0.27319	0.39235	13
84	-0.27747	0.39071	0.40883	0.34133	-0.27755	0.39068	13
85	-0.27623	0.39118	0.41000	0.34090	-0.27632	0.39115	13
86	-0.26404	0.39588	0.42165	0.33664	-0.26412	0.39584	13
87	-0.40148	0.34403	0.29711	0.38319	-0.40152	0.34402	9
88	-0.40609	0.34234	0.29293	0.38479	-0.40613	0.34232	8
89	-0.36216	0.35862	0.33310	0.36953	-0.36222	0.35859	7
90	-0.36807	0.35641	0.32748	0.37165	-0.36812	0.35639	7
91	-0.47842	0.31618	0.22478	0.41107	-0.47844	0.31617	5
92	-0.44448	0.32835	0.25570	0.39909	-0.44451	0.32834	5
93	-0.35486	0.36135	0.33824	0.36759	-0.35492	0.36132	5
94	-0.34368	0.36554	0.34883	0.36361	-0.34374	0.36552	4
95	-0.34943	0.36338	0.34375	0.36552	-0.34949	0.36336	4
96	-0.36689	0.35685	0.32731	0.37172	-0.36695	0.35683	4
97	-0.36935	0.35593	0.32526	0.37249	-0.36941	0.35591	4
98	-0.21705	0.41409	0.46665	0.32037	-0.21715	0.41405	3
99	-0.32870	0.37119	0.37033	0.35557	-0.32873	0.37118	1
100	-0.58704	0.27859	0.14587	0.44201	-0.58696	0.27862	0

Chapter 4: Decomposing the inverse land size-yield relationship

4.1 Introduction

Since it was first observed by Chayanov (1926), economists have struggled to explain the inverse relationship between land size and yield. In its starkest terms, this persistent empirical regularity suggests that the average productivity of cropped land decreases as more land is brought under cultivation. If average productivity of land actually declines as farm size increases, *ceteris paribus,* it could imply that redistributional land reform intended to promote egalitarian goals may also be associated with increased economic efficiency. There are, however, economic reasons for doubt. Perhaps the most important historically is the fundamental belief by many economists in constant-returns-to-scale technologies. And, when properly understood, this belief is hard to controvert. Put simply, it says that if something can be done once, it should be possible to replicate it *in the long run,* that is, when all possible factors of production are freely variable. But few, if any, data sets are complete enough to incorporate information on all possible factors of production. Accordingly, attention has been focused on determining whether the relationship observed by Chayanov (1926) might be an artifact of the way empirical data on land productivity have been measured and collected. In particular, significant effort has been devoted to verifying whether the observed relationship may reflect differences in the rates at which both observed and unobserved inputs (for example, soil quality, human capital) are utilized and measured.

This study does not pretend to resolve the inverse land-size-yield paradox. Instead its primary contribution lies in the observation that yield is, formally speaking, a partial productivity measure, and as such is amenable to decomposition methods developed to analyze productivity difference across economic units. This study, relying on developments in the efficiency measurement and productivity accounting literatures, develops a productivity decomposition for examining the determinants of yield. The procedure does not rely on specific assumptions on returns to scale and accommodates the presence of multiple outputs in the production process. The method decomposes a yield index into six components: (1) efficiency differences, (2) soil-quality differences, (3) a land-size component, (4) a variable-inputs component, (5) a capital-inputs component, and (6) an output-mix component. The decomposition is applied empirically to a sample of Kenyan household farmers.

The rest of this essay is divided in four sections. The next section discusses the methodology, and the third presents the data. The fourth section explains the empirical results, and the last section concludes.

4.2 Methodology

The present methodology is grounded in the observation that yield is a partial productivity measure. Therefore, once yield is converted into index form by comparing it to some base-level yield, simple methods rooted in the theory of index numbers and productivity accounting (Kumar and Russell, 2002) can be used to analyze it exactly as any other partial productivity measure.

Empirical analyses on the inverse land size-yield relationship frequently do not recognize that yield is a partial productivity measure. Lacking that recognition, they cannot accommodate appropriately effects arising from the contemporaneous variation of multiple outputs and multiple inputs. The present contribution adapts methods developed by Färe et al. (1994) and Kumar and Russell (2002) to accommodate variable returns to scale, multiple outputs[1], and a number of production aggregates higher than considered in previous studies.

4.2.1 Multi-output technology

Let $\mathbf{y} \in \mathbb{R}_+^S$ denote a vector of outputs, $\mathbf{x} \in \mathbb{R}_+^U$ denote the variable inputs, $l \in \mathbb{R}_+$ denote land area, $\mathbf{c} \in \mathbb{R}_+^D$ denote soil characteristics, and $\mathbf{b} \in \mathbb{R}_+^R$ denote capital inputs. The multi-output technology set $T \subset \mathbb{R}_+^{U+1+D+R+S}$ is defined:

$$T = \left\{ (\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) \in \mathbb{R}_+^{U+1+D+R+S} : (\mathbf{x}, l, \mathbf{c}, \mathbf{b}) \, \text{can be used to produce} \, \mathbf{y} \right\}.$$

We assume that T satisfies:

A.1: T is nonempty and closed;

A.2: Weak disposability of outputs, that is, $(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) \in T \Longrightarrow (\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \lambda\mathbf{y}) \in T, 0 < \lambda < 1$.

[1]Professor Dr. Richard E. Just suggested the necessity to include more outputs and Professor Dr. Robert G. Chambers suggested this method.

Define the Farrell output-oriented measure of technical efficiency, $E : \mathbb{R}^U_+ \times \mathbb{R}_+ \times \mathbb{R}^D_+ \times \mathbb{R}^R_+ \times \mathbb{R}^S_+ \to \mathbb{R}_+$, by:

$$E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) \equiv \max\{\lambda > 0 \colon (\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \lambda\mathbf{y}) \in T\} \tag{4.1}$$

if there exists $\lambda > 0$ such that $(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \lambda\mathbf{y}) \in T$ and 0 otherwise, and where, by construction, $E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y})$ is positively homogeneous of degree -1 in $\mathbf{y}$. By weak disposability of outputs,

$$E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) \geq 1 \Leftrightarrow (\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) \in T \tag{4.2}$$

so that $E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y})$ is a complete function representation of T. In this multi-output technology, because our interest is in yield-based measures, it proves analytically convenient to partition $\mathbf{y}$ as

$$\mathbf{y} = \left(\mathbf{y}^N, y^M\right)$$

where $\mathbf{y}^N \in \mathbb{R}^{S-1}_+$ and $y^M \in \mathbb{R}_+$. By homogeneity:

$$E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}^N, y^M) = \frac{E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \frac{\mathbf{y}^N}{y^M}, 1)}{y^M}, y^M > 0. \tag{4.3}$$

4.2.2 Yield-index decomposition

Subscripts will be used to denote different farm units. Using 0 to denote the base unit, the yield index for output y^M for unit 1 is defined by

$$\frac{y_1^M/l_1}{y_0^M/l_0}.$$

Using (4.3) yields:

$$\frac{y_1^M/l_1}{y_0^M/l_0} = \frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)/l_1}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)/l_0} \frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \mathbf{y}_0^N, y_0^M)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \mathbf{y}_1^N, y_1^M)}. \tag{4.4}$$

Expression (4.4) naturally breaks into two parts. One,

$$\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \mathbf{y}_0^N, y_0^M)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \mathbf{y}_1^N, y_1^M)},$$

is a measure of relative efficiency between units 0 and 1. The other,

$$\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)/l_1}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)/l_0}, \tag{4.5}$$

is interpretable as an index of "maximum yields" for output y^M. By (4.2), the maximum amount of y^M obtainable given fixed levels of $\left(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}^N\right)$ is

$$\max\left\{y^M : E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) \geq 1\right\},$$

while applying (4.3) gives

$$\begin{aligned} \max\left\{y^M : E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) \geq 1\right\} &= \max\left\{y^M : E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \frac{\mathbf{y}^N}{y^M}, 1) \geq y^M\right\} \\ &= E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \frac{\mathbf{y}^N}{y^M}, 1), \end{aligned}$$

which is interpretable as the maximum amount of output y^M that can be produced, maintaining the fixed output mix, $\frac{\mathbf{y}^N}{y^M}$, for a given amount of $(\mathbf{x}, l, \mathbf{c}, \mathbf{b})$. Dividing by l converts this maximal output into yield terms.

Our conceptual goal is to decompose (4.5) into different components so that one can ascertain the role that differing values of $\left(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \frac{\mathbf{y}^N}{y^M}\right)$ play in determining the overall yield index. By its indicial nature, the observed level of (4.5) depends upon the choice of both 0 (the base unit) and 1 (the unit being compared to 0). That, in turn, means that it is possible to decompose indices such as (4.5) in multiple ways. Illustrating the basic problem using the notation in (4.5) is very burdensome notationally. Therefore, we relegate the details of the actual decomposition to a technical appendix, and content ourselves here with an illustration of the generic problem of creating an index using a function of two variables, $g(u,v)$, so that the index to be decomposed is

$$\frac{g(u_1,v_1)}{g(u_0,v_0)}.$$

Note that we can rewrite this after a simple manipulation as

$$\frac{g(u_1,v_1)}{g(u_0,v_0)} = \frac{g(u_1,v_1)}{g(u_0,v_1)}\frac{g(u_0,v_1)}{g(u_0,v_0)},$$

and one can interpret $\frac{g(u_1,v_1)}{g(u_0,v_1)}$ as an index measuring the effect that variation in u has on the overall index holding v constant and $\frac{g(u_0,v_1)}{g(u_0,v_0)}$ as an index measuring the effect of variation in v holding u constant. However, it's equally possible to write

$$\frac{g(u_1,v_1)}{g(u_0,v_0)} = \frac{g(u_1,v_1)}{g(u_1,v_0)}\frac{g(u_1,v_0)}{g(u_0,v_0)}.$$

Now $\frac{g(u_1,v_0)}{g(u_0,v_0)}$ is the index measuring the effect that variation in u has on the overall index holding v constant, and $\frac{g(u_1,v_1)}{g(u_1,v_0)}$ as measuring the effect of variation in v holding u constant. It's well known that unless g satisfies a specific separability criterion, this arbitrariness in the decomposition is unavoidable in constructing an index number Kumar and Russell (2002); Henderson and Russell (2005). A workable alternative that avoids this arbitrariness is to follow Caves et al. (1982), Färe et al. (1994), Kumar and Russell (2002), and Henderson and Russell (2005) and rely on a "Fisher ideal index" that takes the geometric average of the two decompositions. That is,

$$\begin{aligned}\frac{g(u_1,v_1)}{g(u_0,v_0)} &= \left(\frac{g(u_1,v_1)}{g(u_0,v_1)}\frac{g(u_0,v_1)}{g(u_0,v_0)}\right)^{\frac{1}{2}}\left(\frac{g(u_1,v_1)}{g(u_1,v_0)}\frac{g(u_1,v_0)}{g(u_0,v_0)}\right)^{\frac{1}{2}}\\ &= \left(\frac{g(u_1,v_1)}{g(u_0,v_1)}\frac{g(u_1,v_0)}{g(u_0,v_0)}\right)^{\frac{1}{2}}\left(\frac{g(u_1,v_1)}{g(u_1,v_0)}\frac{g(u_0,v_1)}{g(u_0,v_0)}\right)^{\frac{1}{2}},\end{aligned}$$

where $\left(\frac{g(u_1,v_1)}{g(u_0,v_1)}\frac{g(u_1,v_0)}{g(u_0,v_0)}\right)^{\frac{1}{2}}$ indexes the effect that variation in u has on the overall index holding v constant, and $\left(\frac{g(u_1,v_1)}{g(u_1,v_0)}\frac{g(u_0,v_1)}{g(u_0,v_0)}\right)^{\frac{1}{2}}$ the effect of of variation in v holding u constant.

We show in an appendix that a straightforward extension of this procedure to our problem results in the following decomposition:

$$\frac{y_1^M/l_1}{y_0^M/l_0} = I^{(\mathbf{x}_1,\mathbf{x}_0)}(l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M})L^{(l_1,l_0)}(\mathbf{x}_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M})$$
$$Q^{(c_1,\mathbf{c}_0)}(\mathbf{x}_0,l_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_1,l_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M})K^{(\mathbf{b}_1,\mathbf{b}_0)}(\mathbf{x}_0,l_0,\mathbf{c}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_1,l_1,\mathbf{c}_1,\frac{\mathbf{y}_1^N}{y_1^M})$$
$$Y^{(\frac{\mathbf{y}_1^N}{y_1^M},\frac{\mathbf{y}_0^N}{y_0^M})}(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_0;\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_1)\,EI, \tag{4.6}$$

where

$$EI = \frac{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_0,\mathbf{y}_0)}{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_1,\mathbf{y}_1)}$$

is the relative efficiency index. $I^{(\mathbf{x}_1,\mathbf{x}_0)}$ measures the effect on the yield index that can be associated with varying variable inputs while holding $\left(l,\mathbf{c},\mathbf{b},\frac{\mathbf{y}^N}{y^M}\right)$. Similarly, $L^{(l_1,l_0)}$ is the component of the yield index that can be attributed solely to varying land size, $Q^{(c_1,\mathbf{c}_0)}$ is the component of the yield index that can be attributed solely to varying soil quality, $K^{(\mathbf{b}_1,\mathbf{b}_0)}$ the component attributable to varying capital inputs, and $Y^{(\frac{\mathbf{y}_1^N}{y_1^M},\frac{\mathbf{y}_0^N}{y_0^M})}$ to varying the output mix. Specific formulae for each component of the decomposition are reported in the appendix.

4.2.3 Implementing the yield-index decomposition

One can obtain empirical estimates of the six components by applying non-parametric linear programming methods without specific assumptions on returns to scale.

The nonparametric DEA approximation to the technology T is as follows:

$$\hat{T} = \{(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) \in \mathbb{R}_+^{U+1+D+R+S} : y_s \leq \sum_{i=1}^{n} \gamma_i y_{is},\, s = 1, \ldots, S;$$

$$x_u \geq \sum_{i=1}^{n} \gamma_i x_{iu},\, u = 1, \ldots, U;$$

$$l \geq \sum_{i=1}^{n} \gamma_i l_i;$$

$$c_d = \sum_{i=1}^{n} \gamma_i c_{id},\, d = 1, \ldots, D;$$

$$b_r = \sum_{i=1}^{n} \gamma_i b_{ir},\, r = 1, \ldots, R;$$

$$\text{for}(\gamma_1, \ldots, \gamma_n) \text{ s.t.} \sum_{i=1}^{n} \gamma_i = 1; \gamma_i \geq 0, i = 1, \ldots, n\} \tag{4.7}$$

where i indexes decision-making units. The set of constraints on $\mathbf{c}$ allows only convex combinations of the soil characteristics in the technology.

The approximation to the function E in the productivity decomposition proposed in this essay is done through a Farrell output efficiency score. The function $E(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y})$ can be calculated as follows:

$$\hat{E}(\mathbf{x}, l, \mathbf{c}, \mathbf{b}, \mathbf{y}) = \max e \in \mathbb{R}_+$$

$$\text{s.t. } e y_s \leq \sum_{i=1}^{n} \lambda_i y_{is},\, s = 1, \ldots, S;$$

$$x_u \geq \sum_{i=1}^{n} \lambda_i x_{iu},\, u = 1, \ldots, U;$$

$$l \geq \sum_{i=1}^{n} \lambda_i l_i;$$

$$c_d = \sum_{i=1}^{n} \lambda_i c_{id},\, d = 1, \ldots, D;$$

$$b_r = \sum_{i=1}^{n} \lambda_i b_{ir},\, r = 1, \ldots, R;$$

$$\text{for}(\lambda_1, \ldots, \lambda_n) \text{ s.t.} \sum_{i=1}^{n} \lambda_i = 1; \lambda_i \geq 0, i = 1, \ldots, n. \tag{4.8}$$

The estimation method relies on the choice of the reference farmer $(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \mathbf{y}_0)$. This choice imposes a different normalization on the productivity estimates. Depending on the normalization, the estimates might appear infeasible in some cases. One can minimize the infeasibility problem deriving from the choice of the reference level by applying a 'lattice' approach (Färe et al., 2004).[2] The reference unit, in this study, is a farmer with very high production potential (highest input levels and average soil characteristics) and lowest realized output:

$$(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \mathbf{y}_0) = (\max\{\mathbf{x}\}_u, \max\{l\}, mean\{\mathbf{c}\}_d, \max\{\mathbf{b}\}_r, \min\{\mathbf{y}\}_s),$$

$$\forall u \in 1, \ldots, U,\, d \in 1, \ldots, D,\, r \in 1, \ldots, R,\, s \in 1, \ldots, S. \tag{4.9}$$

[2]If strong disposability were assumed on the soil characteristics, all the linear programs would be feasible with this reference level.

4.2.4 Testing for the inverse land size-yield relationship

The inverse land size-yield relationship suggests that an increase in land size causes a decrease in average product of land normalized on the reference level chosen. In terms of the decomposition, testing for the inverse land size-yield relationship is equivalent to testing:

$$\frac{\partial}{\partial l_i}\left(L^{(l_i,l_0)}(\mathbf{x}_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_i,\mathbf{c}_i,\mathbf{b}_i,\frac{\mathbf{y}_i^N}{y_i^M})\right) < 0. \tag{4.10}$$

One may test this hypothesis through robust weighted least squares regression of the average product $L^{(l_i,l_0)}(\mathbf{x}_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_i,\mathbf{c}_i,\mathbf{b}_i,\frac{\mathbf{y}_i^N}{y_i^M})$ on land size l_i and on a quadratic term of land size l_i^2:

$$L^{(l_i,l_0)} = \zeta_L + \alpha_L l_i + \beta_L l_i^2 + \eta_{Li}. \tag{4.11}$$

If a negative relationship is present, this is evidence of a pure inverse land size-yield relationship.

Even if an inverse relationship exists in (4.11), it is possible that different land average products of interest to the present study are not inversely related to land size. One of the hypotheses of this essay is that long-term soil-quality productivity is heterogeneous. One may consider whether this heterogeneity causes a negative relationship (between a quality-adjusted average land product and land size) through the following robust regression:

$$L^{(l_i,l_0)}Q^{(\mathbf{c}_i,\mathbf{c}_0)} = \zeta_{LQ} + \alpha_{LQ} l_i + \beta_{LQ} l_i^2 + \eta_{LQi}.$$

It is of interest to test also whether short-term productivity-increasing practices (such as usage of fertilizers, or of hand hoes) change the negative relationship of the average land product with land size through the following robust regression:

$$L^{(l_i,l_0)}I^{(\mathbf{x}_i,\mathbf{x}_0)} = \zeta_{LI} + \alpha_{LI}l_i + \beta_{LI}l_i^2 + \eta_{LIi}.$$

In addition, one may test if an average land product, adjusted both for short-term productivity-increasing practices and long-term soil-quality, is negatively correlated to land size through the following robust regression:

$$L^{(l_i,l_0)}Q^{(\mathbf{c}_i,\mathbf{c}_0)}I^{(\mathbf{x}_i,\mathbf{x}_0)} = \zeta_{LQI} + \alpha_{LQI}l_i + \beta_{LQI}l_i^2 + \eta_{LQIi}.$$

On the other hand one may consider whether capital inputs change the negative relationship of average product of land with land size through the following robust regression:

$$L^{(l_i,l_0)}K^{(\mathbf{b}_i,\mathbf{b}_0)} = \zeta_{LK} + \alpha_{LK}l_i + \beta_{LK}l_i^2 + \eta_{LKi}.$$

A further robust regression of interest relates an average land product, adjusted for long-term soil-quality and capital inputs, with land size:

$$L^{(l_i,l_0)}Q^{(\mathbf{c}_i,\mathbf{c}_0)}K^{(\mathbf{b}_i,\mathbf{b}_0)} = \zeta_{LQK} + \alpha_{LQK}l_i + \beta_{LQK}l_i^2 + \eta_{LQKi}.$$

Whether a maximal average producible output index that excludes changes in output mix is correlated with land size can be tested through the following robust regression:

$$L^{(l_i,l_0)}Q^{(\mathbf{c}_i,\mathbf{c}_0)}K^{(\mathbf{b}_i,\mathbf{b}_0)}I^{(\mathbf{x}_i,\mathbf{x}_0)} = \zeta_{LQKI} + \alpha_{LQKI}l_i + \beta_{LQKI}l_i^2 + \eta_{LQKIi}.$$

Finally, one may test whether a maximum yield index measure—that includes appropriately the effect of outputs' diversification, but excludes inefficiency—is correlated with land size by performing the following robust regression:

$$L^{(l_i,l_0)}Q^{(\mathbf{c}_i,\mathbf{c}_0)}K^{(\mathbf{b}_i,\mathbf{b}_0)}I^{(\mathbf{x}_i,\mathbf{x}_0)}Y^{(\frac{\mathbf{y}_i^N}{y_i^M},\frac{\mathbf{y}_0^N}{y_0^M})} = \zeta_{LQKIY} + \alpha_{LQKIY}l_i + \beta_{LQKIY}l_i^2 + \eta_{LQKIYi}.$$

4.3 Data

The empirical data in this study are a random sample of households from 61 sublocations in Kenya. Data have been collected during early 2007 and refer to preceding short and long growing seasons. The survey is named "Research on Poverty, Environment and Agricultural Technologies (REPEAT): Panel studies in Africa". Survey data are obtained from the National Graduate Institute for Policy Studies (21st century Center of Excellence Program) in Japan. The present cross-section sample is composed of 590 household farmers of which only 452 have data on soil quality.[3] The decomposition proposed in this study requires strictly positive maize yield. Maize yield is calculated in kilograms of dry maize equivalent per acre. Only 443 households have a positive production of the main staple dry maize satisfying this condition. To ensure comparability of results among household types, only these households are included in the final sample.

[3]More households are included in the original sampling scheme but are left out of the analysis because have incomplete entries. No significant bias seems to be introduced in this sample reduction step in the output variables. Marginal effects on the probability (in a Probit regression) of being present in the final sample are insignificant for the outputs at 1%. In the case of the inputs, the resulting sample might suffer some bias: farmers who own a lower number of cows, who spend more in hired labor, use more fertilizers, and employ lower amount of seeds are observed more often. These apparent biases are not critical for the methodological results of this paper. They should however be considered when evaluating the representativity of our Kenyan sample. No soil-quality chemical quantitative characteristics in household farmers panel surveys are available in developing countries until the present moment.

In this analysis 9 inputs and 3 outputs are used to represent the technology.[4] Table 4.1 shows input and output summary statistics for the households. Main agricultural output is harvested maize which is totally rain-fed. All present farmers produce maize. Other representative outputs are milk and off-farm income. Apart from land area, the measured variable inputs are seeds, fertilizers, cost of temporary hired workers, and hand hoes. The label 'capital inputs' is used to design the human labor capital, and animal capital of the households. These include household members, adjusted for age and educational level, and number of dairy cows.[5]

[4]Other representations of the technology, for example with a finer subdivision of fertilizers into organic and inorganic, or diversification of maize output into dry and green, or differentiation among household members working off-farm and on-farm, have been tried but results do not change qualitatively.

[5]Faithful to the human capital approach pioneered by Jorgenson and Griliches (1967), on-farm labor input is adjusted for differences in quality, both due to education and age. To adjust labor for education, estimates of the impact of education on agricultural productivity in Kenya from the analysis by Husbands et al. (1996) are used. Given an 85% probability that a primary school-completer household head increases household profitability by 40%, an expected return equal among years of primary education (8 years) obtains a 4.25% average increase in productivity per additional year of any level of education. It is possible that different methods of partitioning the increase in productivity would have achieved a different result. But this topic is not the focus of the present contribution. The same is done for households with a primary school completer, who is not the household head (increase by 29.75%). All of the family members living at the household that are above three years old are included. Most kids start completing tasks in the family at this age. To adjust labor for age, number of children who, by United Nations Children Fund classification, are considered those less than 15 years old, are divided in half. Finally we add the number of members working off-farm.

Data on physical characteristics of land in the largest maize plot of each household are available for mid-2003. The present analysis focuses on critical measures of soil structure and soil texture. The critical characteristic for soil structure is carbon, while for soil texture is clay. These variables indicate the long-term soil quality available to the households. Soil carbon is modifiable only in the long-run, while soil clay is hardly changeable.

Agriculture is mainly manual. Hand hoes are common: the median is 4 hand hoes per household. Median household annual off-farm earnings from all sources are 41100 Kenyan Shillings ($587 in February 2007 dollars). Households at the 25^{th} percentile of the off-farm income distribution earned the equivalent of about $163. The median household head education level is primary. Median application of fertilizers on maize fields is 375 kgs per family. Median household maize harvest is 765 kgs per family. Median household size is 4.5 people, who live at home, plus another family member who works off-farm. Half of the households have at most 1 dairy cow, which may also be used as draft power.

4.4 Results

Figure 4.1 depicts visually one of the main results of this study:[6] the relationship between land area, and the productivity components of land size $L^{(l_i,l_0)}$ and soil quality $Q^{(\mathbf{c}_i,\mathbf{c}_0)}$. The average product of land $L^{(l_i,l_0)}$ is negatively correlated to land acreage, signaling a decreasing average product of land for an increasing area.

[6]Due to infeasibilities in the linear programs under convexity of the soil-characteristics requirement set, only 401 units have complete results.

This negative correlation is evidence of diminishing marginal productivity of land, and it favors the hypothesis of an inverse land size-yield relationship.

Figure 4.1 also pictures the relationship between the soil-quality component and land area. The quality component is more variable among families with medium-size plots than among families with either very small or very large plots, which are of better quality. Farmers with very small plots (less than 0.4 acres) seem to be constrained to produce the staple maize on very good quality plots to obtain a subsistence level of production.

Farmers with medium-size plots need not plant a staple, such as maize, in their best-quality plots to obtain a subsistence level of production. Moreover, if farmers plant more than 2 acres, it is interesting to notice that the variability in soil-quality component decreases towards better quality. Farmers with extensions larger than 2 acres apparently seek obtaining (possibly because of specialization) plentiful maize harvests by planting on good-quality soil.

Figure 4.2 presents a matrix of all six yield components. The graphs on the diagonal are histograms that show the distributions of each component: efficiency, land size, quality, capital inputs, variable inputs, and outputs mix. The graphs outside the diagonal represent each component against the others.

The first element on the diagonal from the upper-left corner evidences how the efficiency scores have a mass around 1. The non-efficient households have a modal peak around 0.4, in the range between 0 and 1.

The second element on the diagonal is the histogram of the average product of land. Most estimates of land average product are concentrated in the lower half of the range, with only very few in the higher half. As can be seen in figure 4.1, low average land product estimates correspond to farmers with large plots, while very large average land product estimates correspond to households with small plots. The negative land size-yield relationship seems to be especially due to these farmers with small plots and with very high average land product.

The third element on the diagonal from the left in figure 4.2 depicts the distribution of the estimated soil-quality components. Even though soil-quality component estimates are more concentrated around and below 1, there are estimates also higher than 1. Numbers higher than 1 reflect the possibility that the average values of reference for soil characteristics are not the values that allow highest producible output under convexity of the soil-characteristics requirement set.

Figure 4.3 plots the soil-quality productivity component estimates against the soil characteristics. The estimates of the soil component are non-monotonic with respect to increases in the soil characteristics. Both soil characteristics show first an increasing, and then a decreasing portion of the soil quality productivity component, for increasing soil characteristics. Estimated soil-quality producible output differences plotted against soil carbon present a steep increase up to around 3% of soil weight, and a smooth decrease thereafter. A decreasing soil-quality component implies congestion of soil carbon after 3%. On the other hand, along the soil-clay distribution, increases and decreases of the soil-quality component are less dramatic and smoother. Nonetheless, soil clay-percentages above 26% appear detrimental to soil quality.

Estimated capital-inputs components are depicted as the fourth plot on the diagonal in figure 4.2. Because of the assumption of strong disposability on these inputs, estimates of the capital-inputs component distribute with increasing probability up to the maximum value of 1.

Non-monotonic distributions, instead, are shown both for the variable-inputs component, the fifth element on the diagonal, and for the output-mix component, the last element on the diagonal. The estimates of the variable-inputs component appear bi-modal, signaling the presence of two distinct groups of households: one group with estimated variable-inputs components less than half the component of the reference farm (mode around 0.2), and one group with estimated components more similar to the reference farm (mode around 0.75). The group with low levels of variable-inputs component hires less labor, and uses lower amount of seeds than the group with higher variable-inputs components.

The estimates of the output-mix component show also a bi-modal distribution: some component estimates are concentrated around 1, while others have lower mode around 0.45. Because the index decomposition is in terms of yield, the farmers who obtain an output mix component around 1 are more specialized in farming than in milk production or in working off-farm. Farmers with an output-mix component lower than 1 are instead substituting more maize harvests with other products, at different degrees.

In this setting, it is expected that farmers diversify production. Figure 4.4 represents the output mixes of the farmers with maize harvest less than 5000 kgs, with less than 10000 liters of milk produced, and less than 600000 KSh of earned off-farm income. As visible in figure 4.4, the farmers with low harvests show diversification in output production, especially in milk production. It is sensible to expect synergies in the production of milk and maize.

In the scatter plots outside the diagonal in figure 4.2 a negative relationship between average product of land and variable-inputs component is present: high estimates of average land product are related to low levels of variable-inputs component. This negative relationship signals that farmers with small plots (who have high average land product) employ less variable inputs than farmers with large plots (who have low average product of land). The same farmers with low usage levels of variable inputs are also somewhat more efficient, compared to farmers who have high usage levels of variable inputs.

The estimates of the variable-inputs component do not appear correlated significantly in any direction to the estimates of the soil-quality components. Different levels of variable-inputs components are roughly possible at any level of quality component.

High values of the quality component appear correlated to high capital-inputs components, to low levels of efficiency, and to low or medium levels of the average land product. In other words, high soil quality is in highly capitalized farmers with relatively large plots. This is evidence, as shown in figure 4.1, of an unequal distribution of soil quality along the land-size distribution.

4.4.1 Test results on the inverse land size-yield relationship

In figures from 4.5 to 4.12 and from 4.17 to 4.18, ten graphs show different productivity aggregates and the results of the fitted robust regressions against land size. The statistical results related to the robust regressions are in table 4.2.

The plot in figure 4.5 represents the average product of land $L^{(l_i,l_0)}$ against land size. There is a strong negative relationship between the average product of land $L^{(l_i,l_0)}$ and land area. The results support hypothesis (4.10) of a negative relationship between yield and land area. This key result for the present study is corroborated by strongly negatively significant coefficients in table 4.2.[7] The results support a convex decay of the average product of land at increases in land area. The convex decay implies fast decreasing marginal productivity of land in this setting. The convexity also reinforces the finding that such an inverse land size-yield relationship may be due to small farmers with very high average products of land.

The graph in figure 4.6 pictures a quality-adjusted average product of land, derived as the multiplication of the quality component $Q^{(\mathbf{c}_i,\mathbf{c}_0)}$ and the average product of land $L^{(l_i,l_0)}$, against land area. The graph shows that even a quality-adjusted average product of land is negatively correlated to land area. Indeed, the higher variability in the long-term soil-quality components in figure 4.1 comes mostly from the farmers in the middle of the land size distribution. In these cases the quality-adjusted average product of land is lower. Even though the precision and the

[7]The result is robust when repeating the calculations with a bias-corrected regression as proposed in Kneip et al. (2015a).

strength of the negative coefficients of the first panel of table 4.2 are weakened, the results still support a convex decay of the quality-adjusted average product of land, when land area increases.

The graph of figure 4.7 shows how land area is related to a variable input-adjusted average product of land, derived as the multiplication of the average product of land $L^{(l_i,l_0)}$ and the variable-inputs component $I^{(\mathbf{x}_i,\mathbf{x}_0)}$. This variable input-adjusted land component looks at the possibility that short-term productivity increasing agricultural practices (such as usage of fertilizers, or of hand hoes) might reduce the intensity of the negative land size-yield relationship.

The coefficients in the quadratic regression of a variable input-adjusted average product of land on land size are in table 4.2. Even though the intensity and the precision of the coefficients are lower than when considering only the average product of land, they maintain a (positively) convex decreasing relationship between an average product of land (adjusted for short-term productivity increasing practices) and land area. Short-term productivity increasing practices have a stronger impact in diluting the inverse land size-yield relationship than long-term soil quality has. This is due to the fact that the contribution of variable inputs to the yield of farmers with small land plots is low compared to the reference farm.

The graph in figure 4.8 looks at the interaction of short-term productivity-increasing practices $I^{(\mathbf{x}_i,\mathbf{x}_0)}$ and long-term soil-quality $Q^{(\mathbf{c}_i,\mathbf{c}_0)}$ with average product of land $L^{(l_i,l_0)}$, against land size. From the graphical results it is possible to see that the inverse land size-yield relationship is further diluted by interacting soil quality and variable inputs with average product of land. The statistical results confirm this intuition in table 4.2. The coefficients that characterize a (positively) convex decreasing relationship remain significant even if they are weaker than when each productivity component ($I^{(\mathbf{x}_i,\mathbf{x}_0)}$ or $Q^{(\mathbf{c}_i,\mathbf{c}_0)}$) is considered alone in interaction with $L^{(l_i,l_0)}$.

The graph in figure 4.9 portrays an average product of land adjusted for capital inputs $K^{(\mathbf{b}_i,\mathbf{b}_0)}$. The (positively) convex decreasing relationship between capital-adjusted average product of land and land size remains strongly significant. The contribution of the capital inputs to yield is rather high for farmers with small plots. Including, additionally to capital inputs $K^{(\mathbf{b}_i,\mathbf{b}_0)}$, long-term soil quality $Q^{(\mathbf{c}_i,\mathbf{c}_0)}$ in figure 4.10 leaves the relationship significant and convex. It is, however, worth noting that the long-term soil-quality component mitigates the inverse relationship further when interacted with capital inputs.

The graph of figure 4.11 displays an index of maximal average producible output ($L^{(l_i,l_0)}\ Q^{(\mathbf{c}_i,\mathbf{c}_0)}\ K^{(\mathbf{b}_i,\mathbf{b}_0)}\ I^{(\mathbf{x}_i,\mathbf{x}_0)}$), which excludes the effects of outputs' diversification, against land size. Despite the relationship is weaker than in previous regression coefficients of table 4.2, there is still a significantly (positively) convex decreasing relationship of $L^{(l_i,l_0)}\ Q^{(\mathbf{c}_i,\mathbf{c}_0)}\ K^{(\mathbf{b}_i,\mathbf{b}_0)}\ I^{(\mathbf{x}_i,\mathbf{x}_0)}$ with land size. From the comparison with other regression coefficients, it is clear that the impact of the variable-inputs component is strong in diminishing the precision of the quadratic term estimate. Indeed, farmers with small plots and high average land product use lower levels of variable inputs than farmers with big plots.

Finally, the plot in figure 4.12 presents a maximum yields measure that accounts for outputs' diversification ($L^{(l_i,l_0)}Q^{(\mathbf{c}_i,\mathbf{c}_0)}\ K^{(\mathbf{b}_i,\mathbf{b}_0)}\ I^{(\mathbf{x}_i,\mathbf{x}_0)}\ Y^{(\frac{\mathbf{y}_i^N}{y_i^M},\frac{\mathbf{y}_0^N}{y_0^M})}$) against land size. The highest values of $L^{(l_i,l_0)}Q^{(\mathbf{c}_i,\mathbf{c}_0)}K^{(\mathbf{b}_i,\mathbf{b}_0)}I^{(\mathbf{x}_i,\mathbf{x}_0)}$, which exclude the output-mix component (figure 4.11), are decreased further when including the output diversification. The numerical results in table 4.2 show a totally insignificant quadratic term, and a weakly significant linear effect. The low significance of the linear effect is jeopardized by the insignificance of the quadratic term. Theory wants that the marginal effect of the lower-order term in this quadratic regression cannot be interpreted if the higher-order term is insignificant. Nonetheless, if one looks at a simple linear relationship, there is still a negatively significant linear correlation. The numerical coefficients are presented in table 4.2.

Considering appropriately outputs' diversification attenuates the decreasing productivity of land by decreasing the maximum yields of farmers with small plots, and thus eliminating the convexity in the relationship. Among farmers with small plots, the estimated outputs components are lower than among farmers with large plots. As can be seen from figures 4.13 to 4.16, farmers with small plots, with low estimated outputs components, are more diversified into non-maize production than farmers with large plots. The productivity effect of output mix, in yield terms, is lower for these farmers who produce more diversified outputs, than for farmers more specialized into maize. Accounting appropriately for the contribution in yield terms of farmers' diversification into other outputs, while controlling for inefficiency, dissipates the convexity of the inverse land size-yield relationship.

In figure 4.17 the efficiency estimates and in figure 4.18 the maize yields are plotted against land area. An inverse (positively) convex relationship of efficiency with land size is apparent. Many farmers with plots less than 1 acre (most with plots below 0.4 acres) and some farmers with plots above 4 acres are completely efficient. An increasing variability of efficiency estimates is visible among farmers with plots between 0.4 and 4 acres. The coefficients in table 4.2 support this conclusion. Similarly to hypothesis (4.10), this result corroborates the inverse land size-yield relationship. As expected, the results in figure 4.18 and the associated numerical coefficients in table 4.2 support a (positively) convex decay of yield when increasing land size.

4.5 Conclusions

The methods presented in this essay reconsider the inverse land size-yield relationship by introducing a decomposition method to analyze its robustness. The proposed method accounts for productivity differences of inputs and outputs without assuming specific returns to scale, production efficiency, nor a parametric form on the technology. The possibility of production inefficiency allows decomposing a ratio of maximum yields and not of observed yields. Absence of specific technological functional form assumptions (apart from piecewise linearity) allows not imposing, a priori, unrealistic properties among inputs and outputs, and specific returns to scale.

The methodology purges out the inefficiency from a yield index, and directly decomposes, under variable returns to scale, the distance among maximum yields into five components: land size, soil-quality characteristics, variable inputs, capital inputs, and output mix. The present study proposes to calculate the six components with nonparametric productivity accounting methods.

The results show, visually and quantitatively, a strongly significant convex inverse relationship between yield and land size. The yield components estimated show a strongly significant (positively) convex inverse relationship of average product of land with land area, which is evidence in support of an inverse land size-yield relationship. By accounting appropriately, in yield terms, for diversification into products different from maize, the convexity of the inverse land size-yield relationship disappears in favor of a linear inverse relationship.

The results imply that in this sample of Kenyan households the long-standing empirical phenomenon of an inverse land size-yield relationship is confirmed. If the implications of this analysis were to be pushed further, farmers with smallest maize plot sizes would be fundamental agents in increasing agricultural production for raising global food needs. These farmers with very small plots use little levels of agricultural inputs in an efficient manner on good-quality soil to produce not only maize but a portfolio of differentiated products. These conclusions are nonetheless only valid for this sample and for these technology assumptions.

Considering the importance of dynamics in soil fertility, this study is moreover only an approximation of the results obtainable if soil-quality household panel data were available. Once these data were to become available, a generalized version of this study would be possible. Such a generalized study would allow disentangling completely the dynamic interplay of long-term soil quality, household choices, and the land size-yield relationship.

4.6 Figures and Tables

Figure 4.1: Land size against productivity components of land size and soil quality

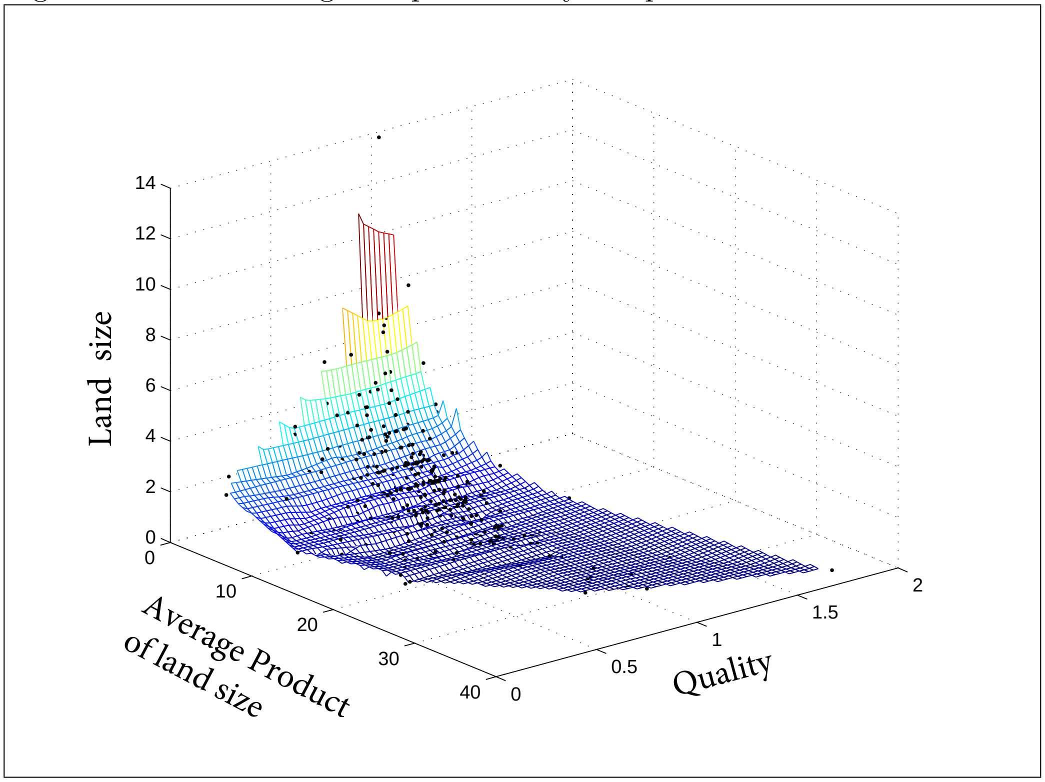


Figure 4.2: Matrix of plots of yield components under variable returns to scale

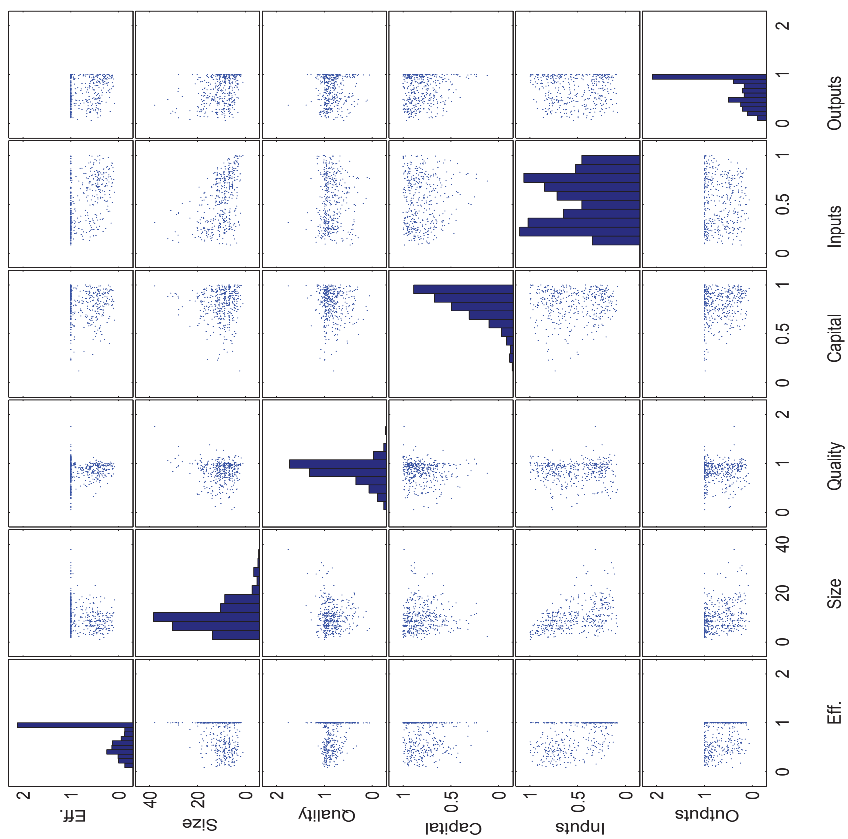


Figure 4.3: Soil-quality component against soil carbon and soil clay

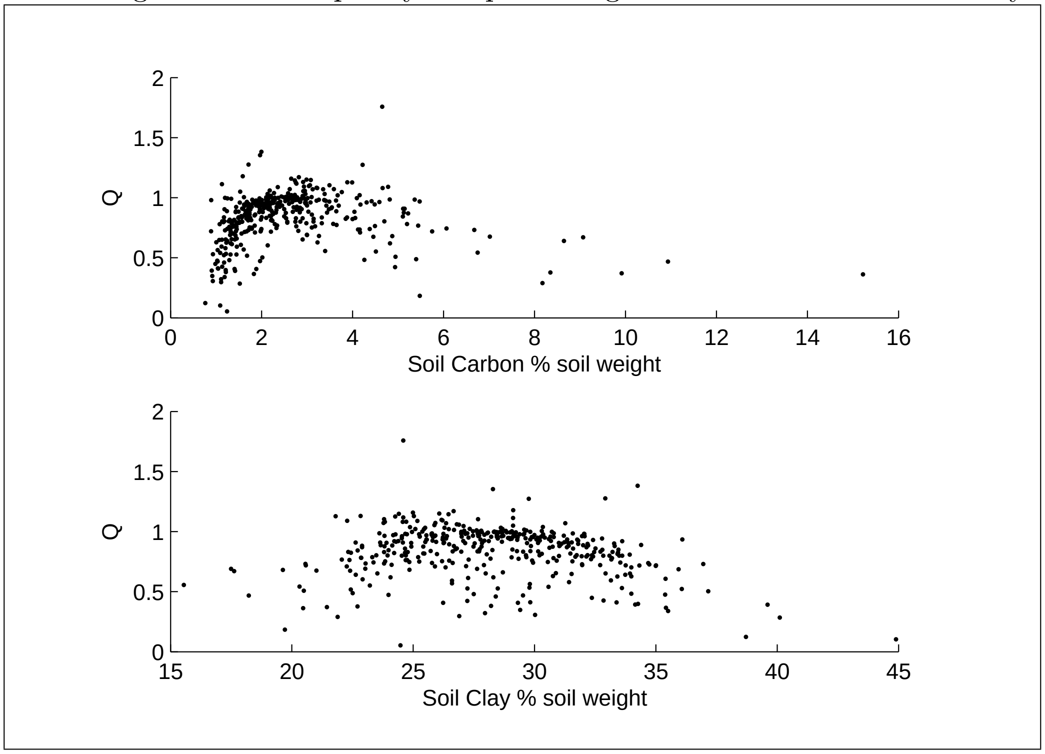


Figure 4.4: Milk production and off-farm income against maize harvest

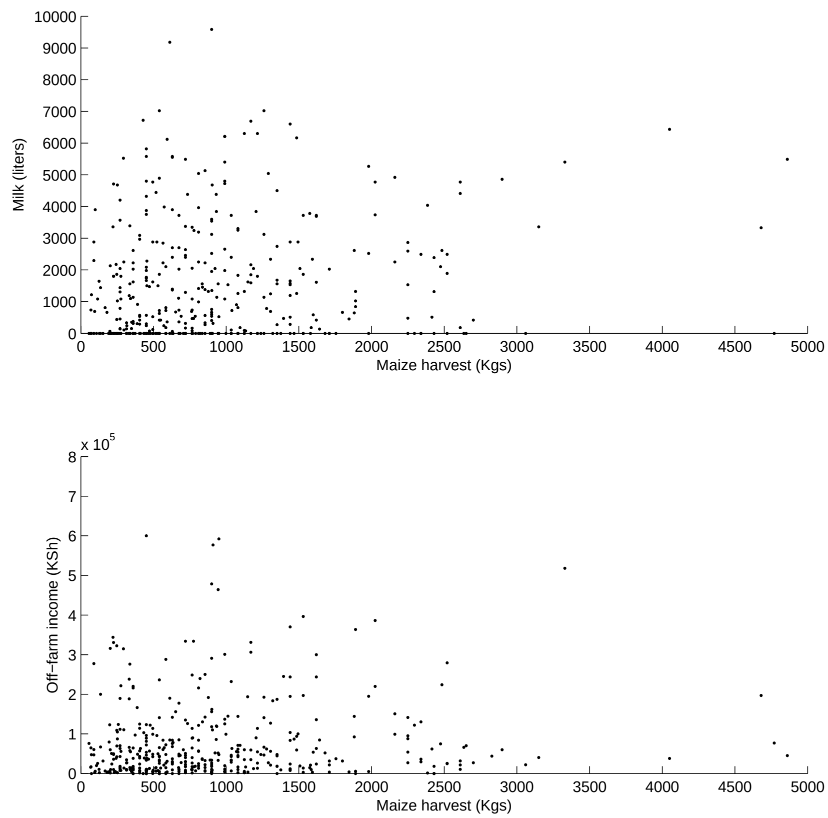
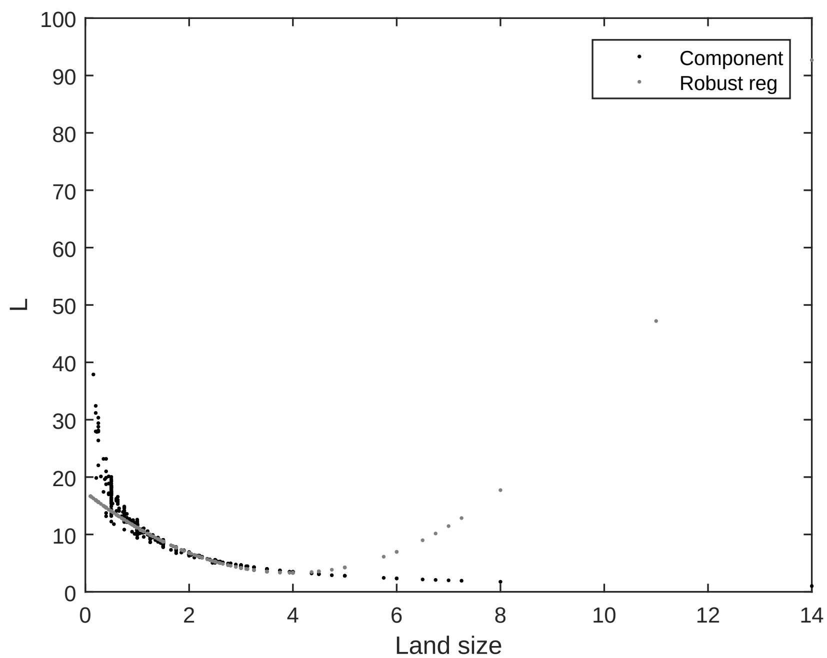
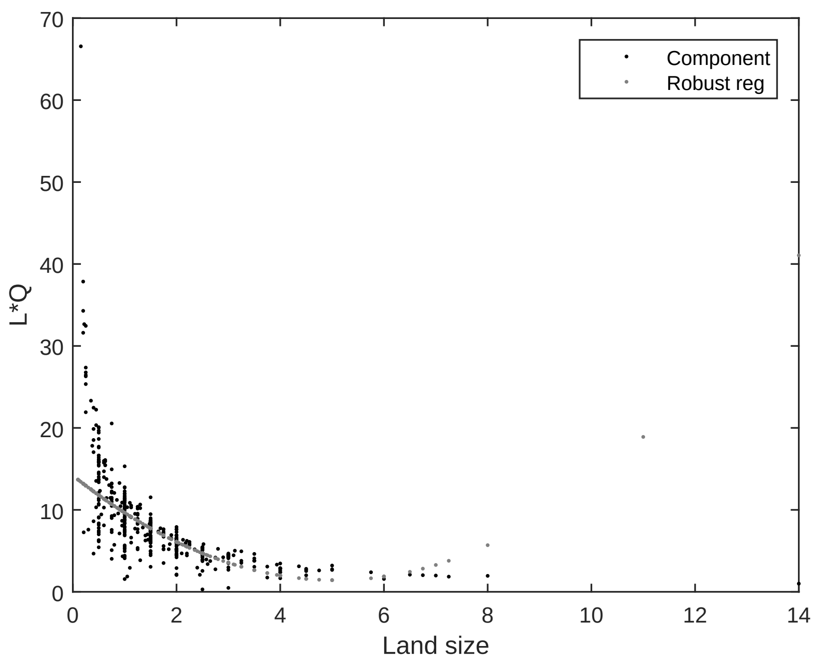


Figure 4.5: Robust quadratic regression of average product of land size on land size under variable returns to scale for a multi-output technology



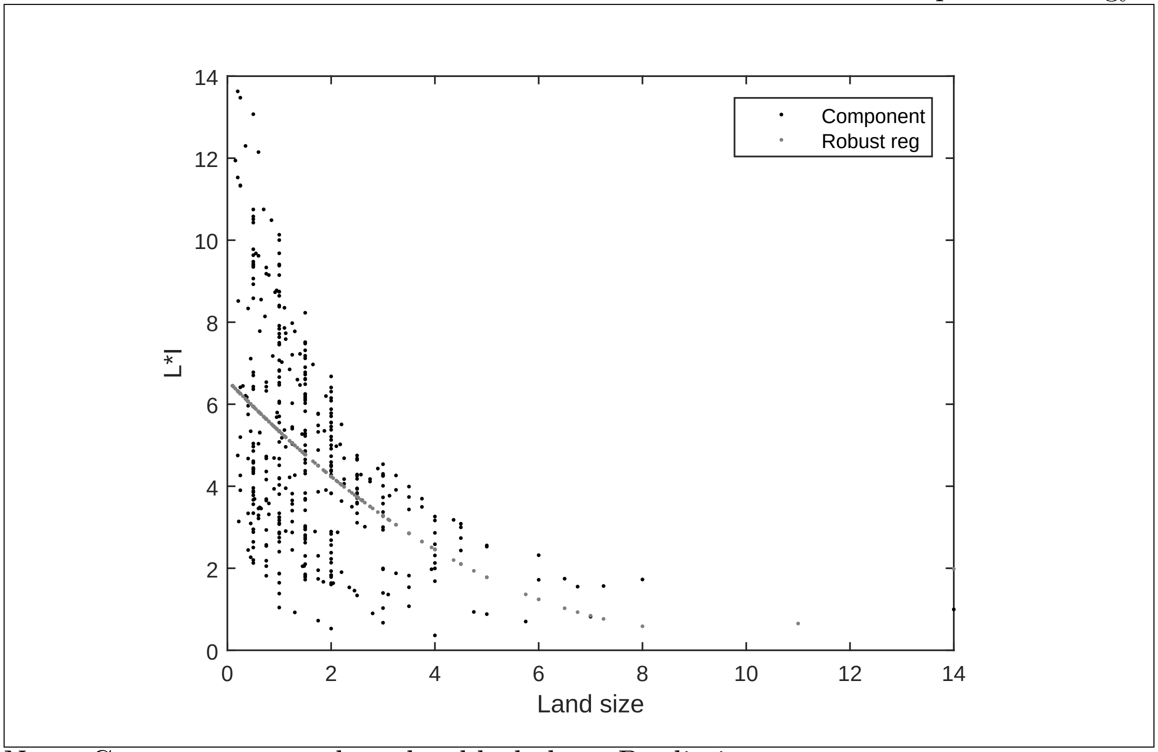
Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.6: Robust quadratic regression of quality-adjusted average product of land size on land size under variable returns to scale for a multi-output technology



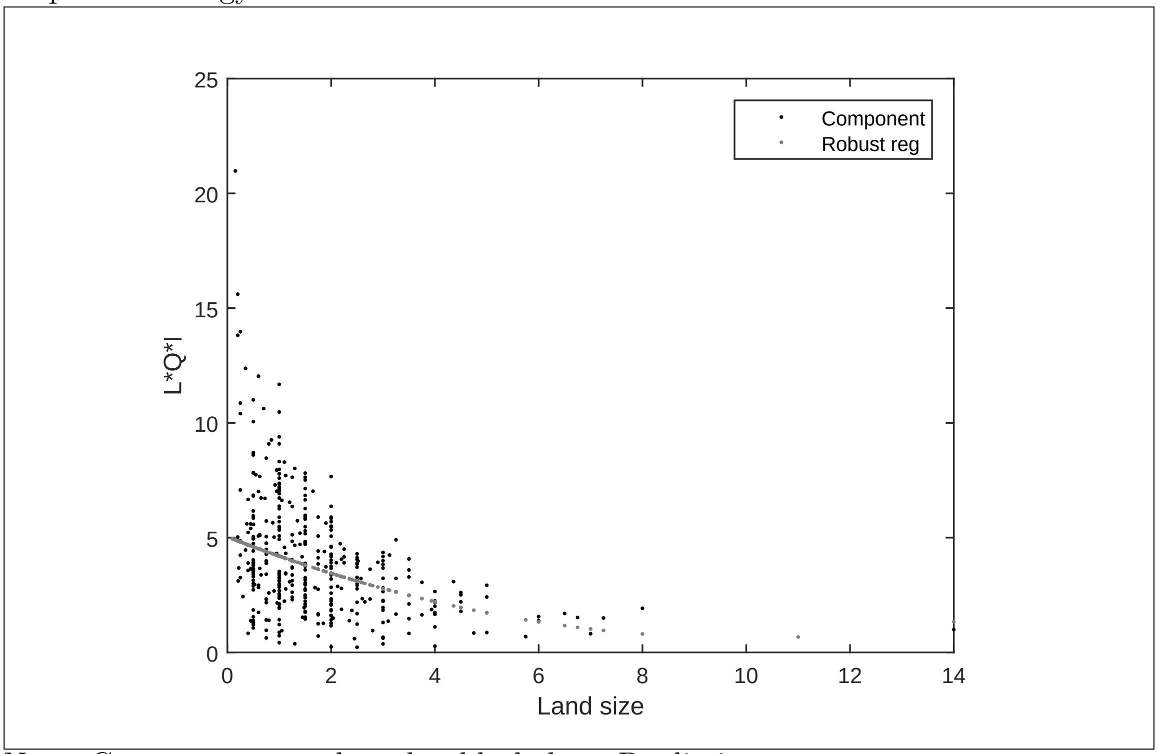
Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.7: Robust quadratic regression of variable inputs-adjusted average product of land size on land size under variable returns to scale for a multi-output technology



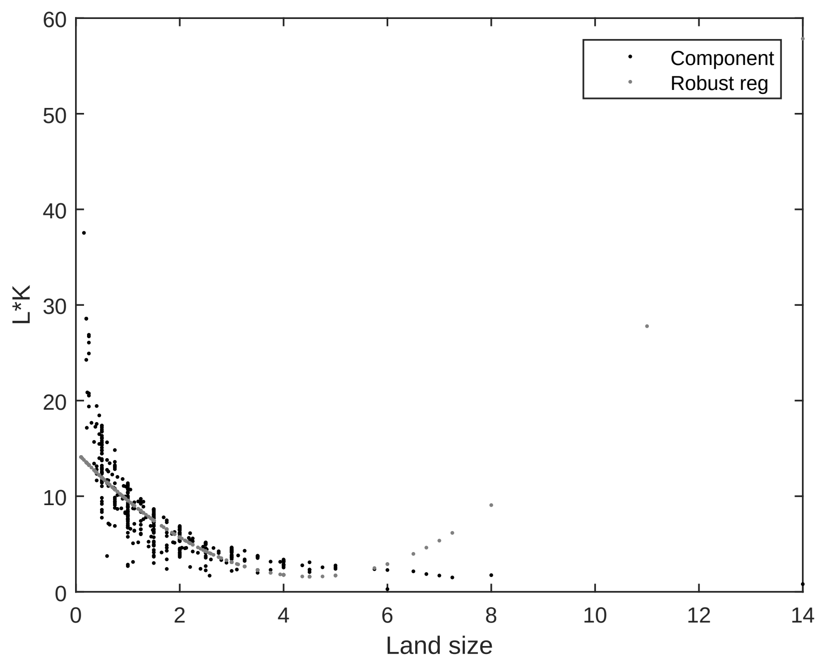
Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.8: Robust quadratic regression of quality and variable inputs-adjusted average product of land size on land size under variable returns to scale for a multi-output technology



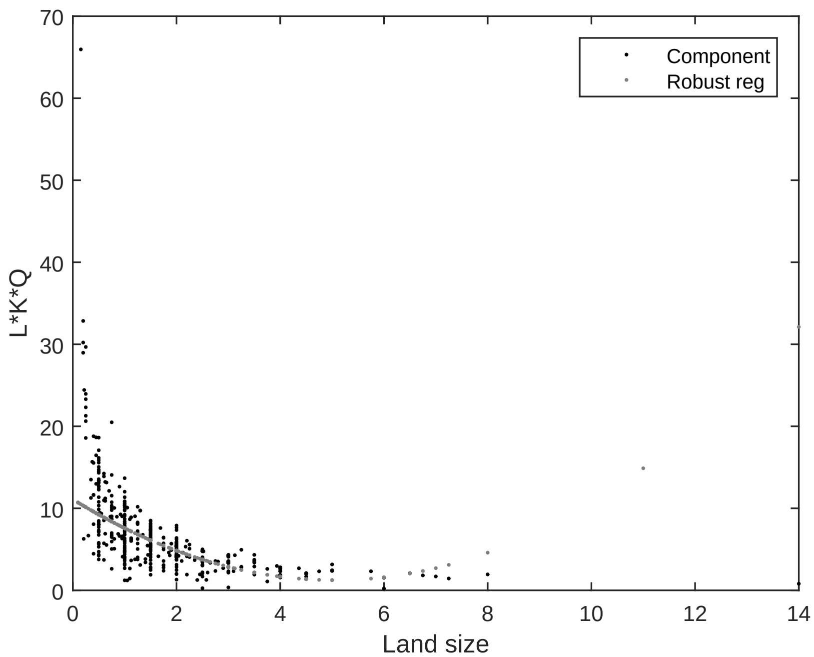
Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.9: Robust quadratic regression of capital inputs-adjusted average product of land size on land size under variable returns to scale for a multi-output technology



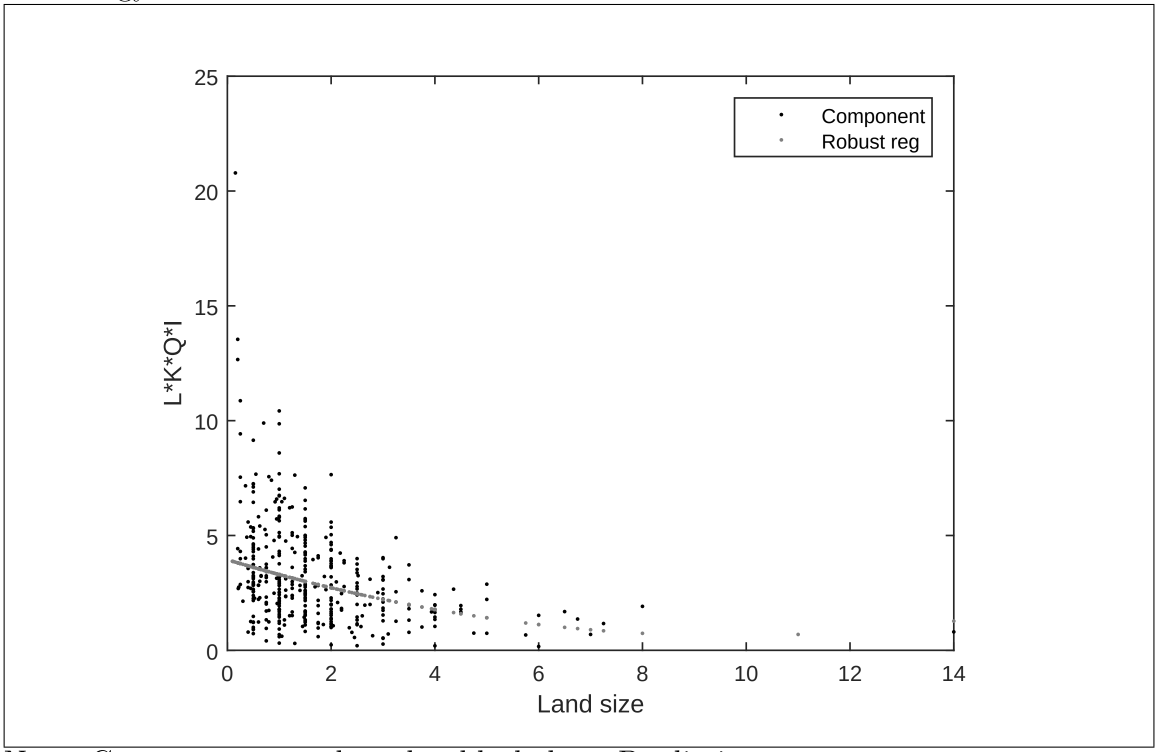
Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.10: Robust quadratic regression of capital inputs and quality-adjusted average product of land size on land size under variable returns to scale for a multi-output technology



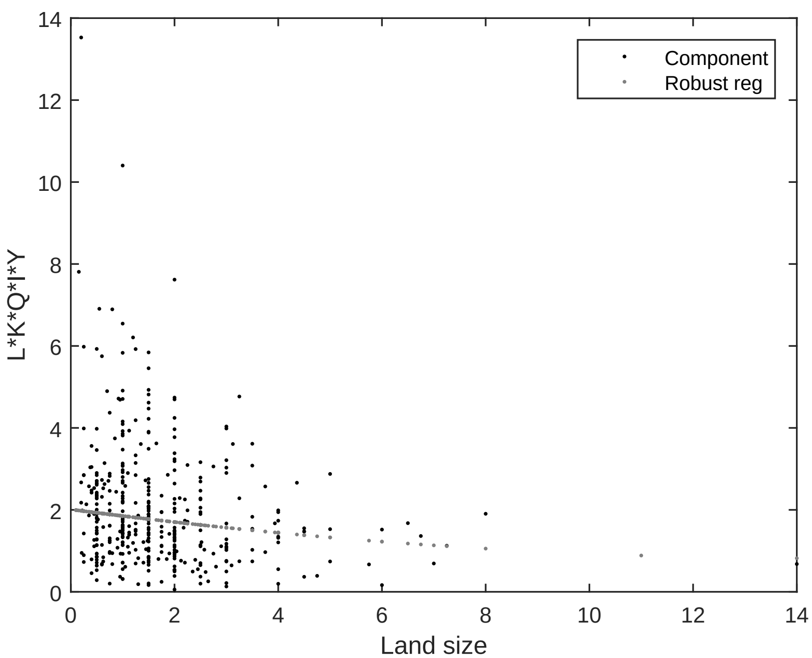
Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.11: Robust quadratic regression of maximal producible output excluding output diversification on land size under variable returns to scale for a multi-output technology



Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.12: Robust quadratic regression of maximum yields on land size under variable returns to scale for a multi-output technology



Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.13: Maize harvest against output-mix component

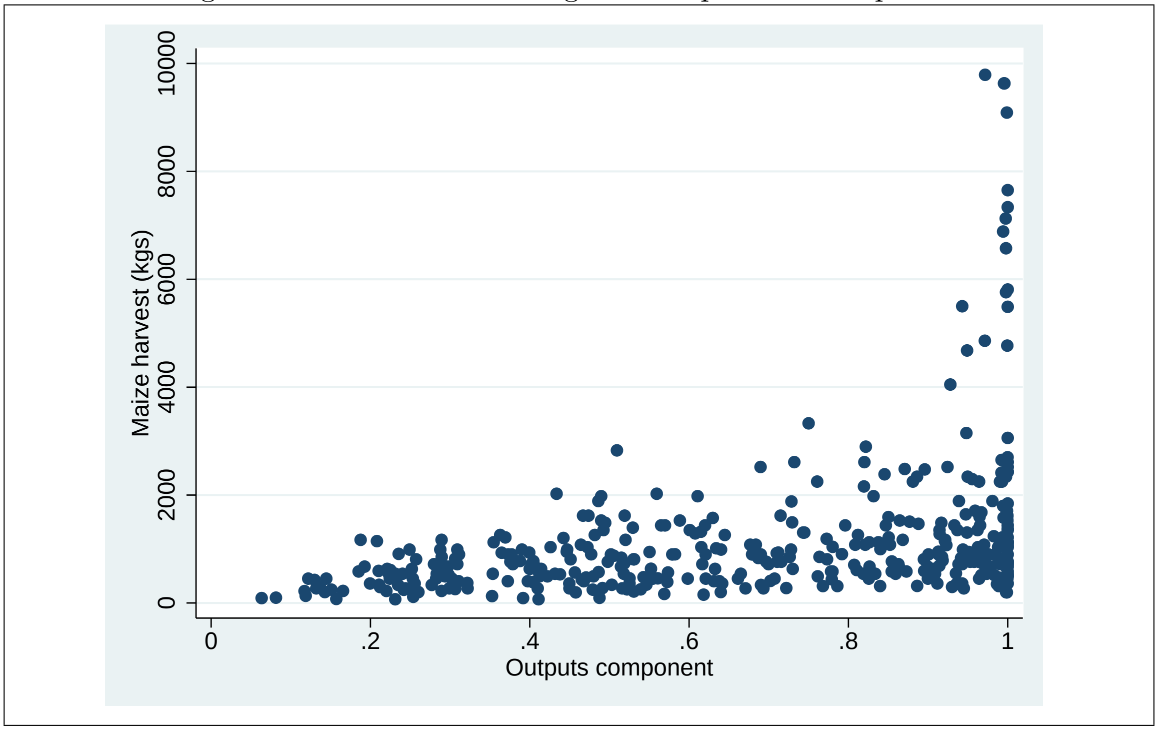


Figure 4.14: Milk production against output-mix component

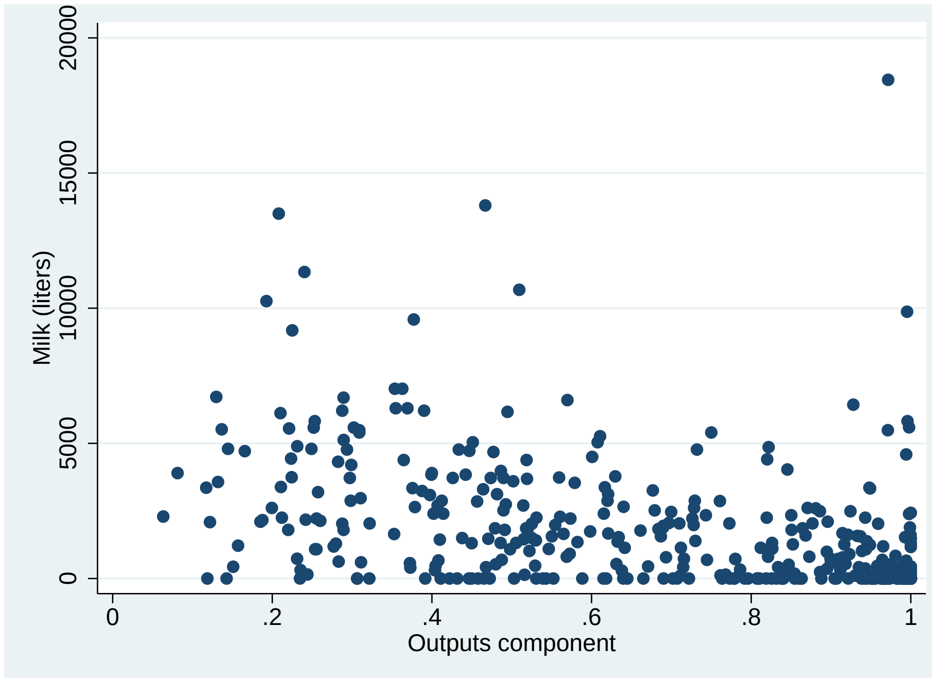


Figure 4.15: Off-farm income against output-mix component

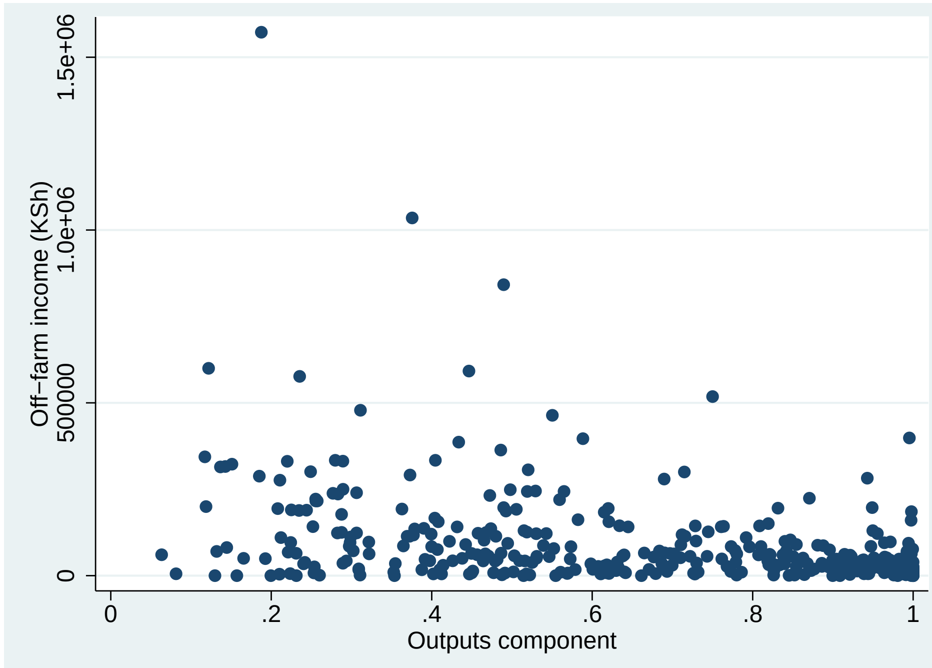


Figure 4.16: Average product of land size against output-mix component

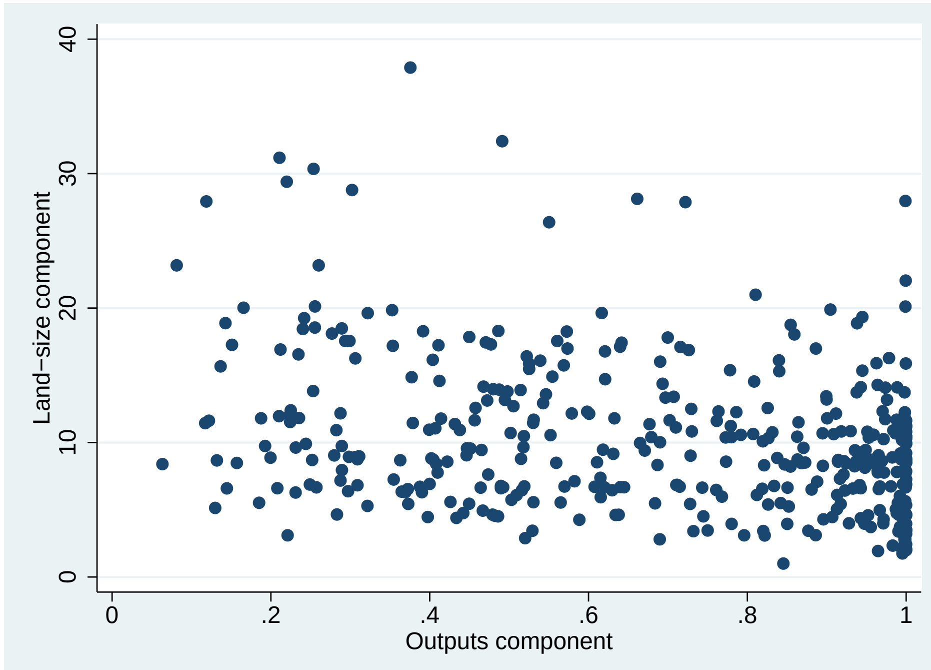
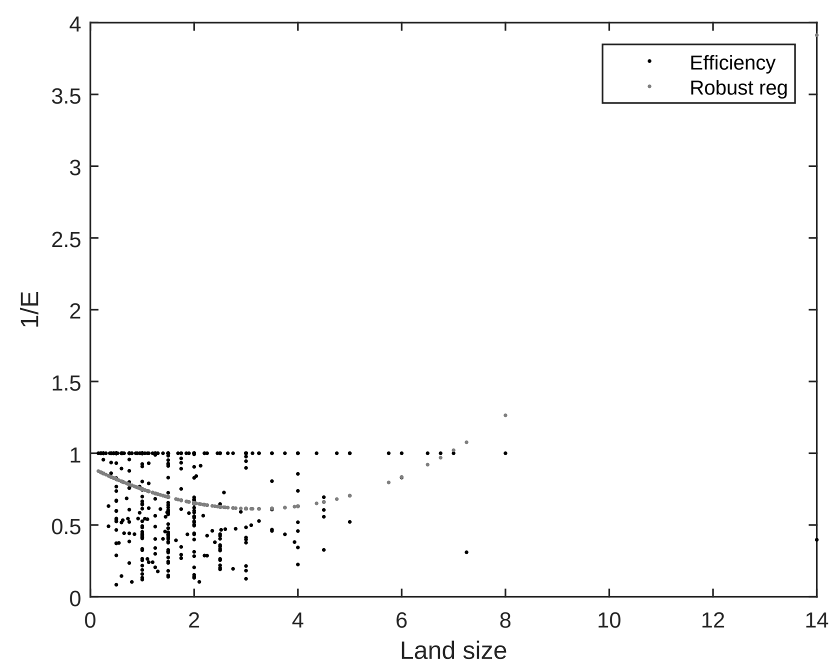
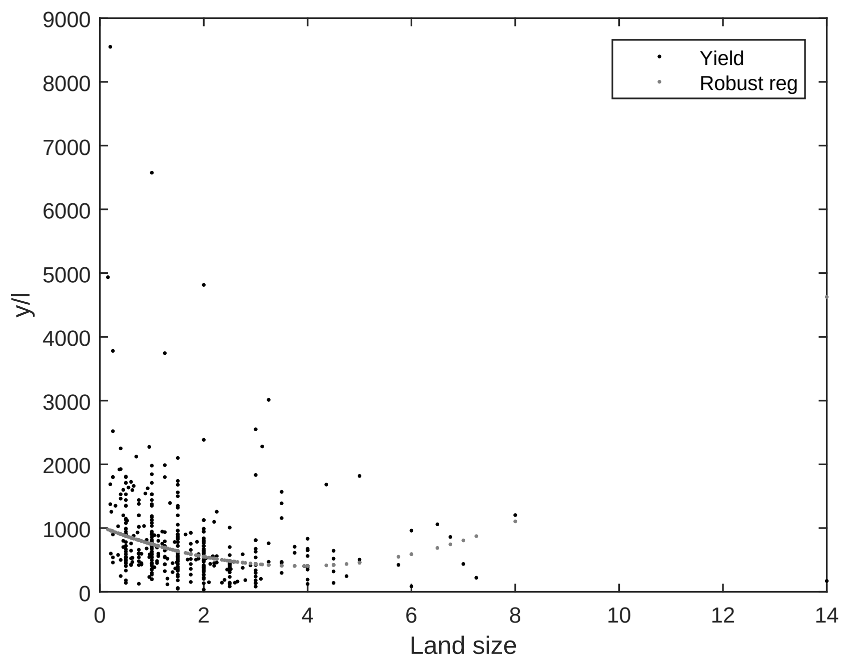


Figure 4.17: Robust quadratic regression of efficiency on land size under variable returns to scale for a multi-output technology



Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Figure 4.18: Robust quadratic regression of yield on land size under variable returns to scale for a multi-output technology



Note: Components are plotted as black dots. Predictions from robust quadratic regression are plotted with light gray dots.

Table 4.1: Summary statistics of inputs, outputs, and soil-quality physical characteristics

Variable	Mean	Std.Dev.	Min	Max
Inputs				
fertilizers (kgs)	800.1	1233.5	0	9000
quantity of seeds (kgs)	13.4	10.9	1	78
number of hand hoes	3.9	2.1	0	15
hired labor (cost in KSh)	2975.3	4940.4	0	48160
land area (acres)	1.6	1.4	0.1	14
members of the family	5.7	2.8	1	17.5
milking cows	1	0.9	0	5
Outputs				
total harvest maize (kgs of dry maize equivalent)	1114.8	1337.9	56.2	9790
milk (liters)	1714.2	2449.1	0	18600
off-farm income (in KSh)	85941	147022.5	0	1572400
Soil-quality physical characteristics				
soil-carbon content (% of soil weight)	2.6	1.5	0.8	15.2
soil-clay content (% of soil weight)	28.3	3.9	15.5	44.9
Observations	443			

Table 4.2: Robust regression coefficient estimates of different yield components under variable returns to scale for a multi-output technology

	Coefficient	SE	T statistic	P Value
Dependent variable: L				
Constant	17.359	0.080	218.092	0.000
l	-7.067	0.056	-126.498	0.000
l^2	0.889	0.006	139.311	0.000
Dependent variable: L*Q				
Constant	14.187	0.240	59.045	0.000
l	-5.036	0.169	-29.864	0.000
l^2	0.497	0.019	25.788	0.000
Dependent variable: L*I				
Constant	6.580	0.234	28.111	0.000
l	-1.311	0.164	-7.979	0.000
l^2	0.070	0.019	3.739	0.000
Dependent variable: L*Q*I				
Constant	5.032	0.209	24.076	0.000
l	-0.882	0.147	-6.010	0.000
l^2	0.044	0.017	2.632	0.009
Dependent variable: L*K				
Constant	14.664	0.190	77.129	0.000
l	-5.745	0.133	-43.050	0.000
l^2	0.631	0.015	41.370	0.000
Dependent variable: L*Q*K				
Constant	11.085	0.260	42.674	0.000
l	-3.897	0.182	-21.373	0.000
l^2	0.386	0.021	18.515	0.000
Dependent variable: L*Q*K*I				
Constant	3.944	0.175	22.526	0.000
l	-0.680	0.123	-5.535	0.000
l^2	0.035	0.014	2.489	0.013
Dependent variable: L*Q*K*I*Y				
Constant	2.010	0.117	17.247	0.000
l	-0.165	0.082	-2.020	0.044
l^2	0.006	0.009	0.613	0.540
Dependent variable: L*Q*K*I*Y				
Constant	1.955	0.089	21.883	0.000
l	-0.119	0.041	-2.887	0.004
Dependent variable: Efficiency				
Constant	0.903	0.031	29.045	0.000
l	-0.181	0.022	-8.308	0.000
l^2	0.028	0.002	11.355	0.000
Dependent variable: Yield				
Constant	1026.35	41.195	24.915	0.000
l	-319.944	28.914	-11.065	0.000
l^2	41.221	3.303	12.479	0.000

Chapter 5: Conclusions

This dissertation extends the theory of production economics with methods to identify separable production aggregates and with methods to test separability. In particular, it focuses on the importance of soils in development economics by quantifying a separable soil-quality measure. The generalization of productivity decomposition methods is another contribution, which relates to productivity theory. An important outcome is the analysis of the role of soils in the debate on the inverse land size-yield productivity relationship.

The first essay quantifies a soil-quality measure by aggregating soil-quantitative characteristics that are assumed separable from other production factors. The measure incorporates the possibility that soil characteristics have negative marginal products. From the application of these methods to the empirical Kenyan case the first essay obtains, under variable returns to scale, a soil-quality measure aggregating soil carbon and soil clay via a separability assumption. This measure has a peak over the observable ranges of soil carbon and soil clay. In other words, negative marginal effects are visible for soil-carbon percentages of soil weight higher than 3 or 4%, depending on the level of soil clay. Moreover, negative marginal products are present for soil-clay percentages higher than 35%, depending on the amount of soil carbon.

The second essay defines a method to test non-parametrically the separability hypothesis assumed in the first essay, without imposing specific assumptions on returns to scale. The essay considers the slow convergence of nonparametric efficiency estimators to obtain two versions of the statistical test. Either version of the test is appropriate depending on the number of inputs and outputs in the technology. The separability assumption on the soil-quality characteristics in the Kenyan case is supported by the test at any reasonable statistical level.

The third essay provides a flexible framework to decompose productivity differences in a multi-input, multi-output technology. The methodology does not rely on parametric assumptions on the technology, nor on specific returns to scale. The estimation of the nonparametric productivity accounting methods, through DEA estimators, yields important results from a policy viewpoint. Under variable returns to scale, a properly developed average product of land size is negatively related to land size in a strictly convex manner. Different aggregate average products of land are defined by successively aggregating estimated productivity components. For example, by multiplying the soil-quality component, one obtains a quality-adjusted average product of land. The inverse land size-yield relationship is still confirmed through quadratic regressions of different aggregate average products of land on land size.

All the methods in this thesis are presented in the case of a radial Farrell output measure. However, these methods can be adapted to other characterizations of the technology. For example, the methodology in the first essay can be extended to consider an output aggregate, instead of an input aggregate. Also the methods in the second essay are adaptable to an output aggregate measure. Finally, the methodologies in the third essay can be rewritten in the input direction if an input index has to be decomposed. Moreover, the methods presented in this thesis can be easily extended to the inclusion of other dimensions or to the case of non-radial input or output measures.

From a general viewpoint, all the methods in this thesis can also be applied to other contexts. For example, the methods presented in the first essay are appropriate tools to analyze climate change or research and development. In the case of climate change, one could measure aggregated climate variables over time. This would identify the changing impact of climatic stochastic inputs on the production process. The present method would obtain directly a characterization of negative marginal effects of climatic stochastic inputs. Among them, flooding effects or heat stress effects could be revealed. In the case of research and development, one could quantify the role of research and development expenditures, from public or private sources, in product innovation.

The methods in the second essay are potentially useful in any production context in which a separability assumption has to be tested.

In the same manner, the methods in the third essay can also be applied to other contexts, such as labor productivity. In this case, labor productivity could be decomposed in contributions due to human capital change or to technological change. These are just some of the possible applications of the general methods in this thesis.

Appendix A: Yield-index decomposition

This study proposes a decomposition of a ratio of maximum yields:

$$\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)/l_1}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)/l_0}, \tag{A.1}$$

It is possible to obtain different decompositions of (A.1). To illustrate, it is possible to first multiply and divide by $E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)$ $E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)$. In this manner one obtains:

$$\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)/l_1}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)/l_0} = \frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)} \frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}$$

$$\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)} \frac{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)/l_1}{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)/l_0} \frac{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}. \tag{A.2}$$

These five ratios on the right-hand side are legitimate index numbers. That is, only one argument changes in every ratio, and every ratio measures relative changes in maximal producible yield due to that varying argument. In particular, the first of the right-hand side terms represents the distance between two maximal producible yields given by a change in output mix. The second of the right-hand side terms evaluates instead the change due to capital inputs. The third of the right-hand side terms measures the change due to soil characteristics. The fourth term considers the change due to land size. Finally, the last of the right-hand side terms quantifies the change due to variable inputs, keeping normalized outputs, capital inputs, soil characteristics, and land size fixed.

The proposed decomposition is not unique. In particular, it is also possible to decompose (A.1) by multiplying and dividing by

$$E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1).$$

This obtains:

$$\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)/l_1}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)/l_0} = \frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)} \frac{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)/l_1}{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)/l_0}$$

$$\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)} \frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)} \frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}. \tag{A.3}$$

Also in this case every right-hand side term represents a proper index. But terms corresponding to changes in same variables in (A.2) and (A.3) are not necessarily the same. For example, the soil-quality component need not be the same in the two decompositions:

$$\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)} \neq \frac{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}. \tag{A.4}$$

This problem is well known in the productivity literature and it is referred to as path dependency. To illustrate the path dependency issue, one can consider figure A.1 where, for exemplifying purposes, only one input x and land size l change. Different paths attribute different measures to changes in x and l. One can determine the paths by changing the variables in different orders. To illustrate, let the comparison be the change between $g(x_1, l_1)$ and $g(x_0, l_0)$ where $g : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$. One can either move from point $g(x_1, l_1)$ to point $g(x_1, l_0)$, and then to point $g(x_0, l_0)$ (first path).

But one can also move from point $g(x_1, l_1)$, to point $g(x_0, l_1)$, and then to point $g(x_0, l_0)$ (second path). The problem of path dependency arises because, as in this example in the picture, the portions of the change from $g(x_1, l_1)$ to $g(x_0, l_0)$ attributed to each component are different depending on the path followed. In figure A.1 only two different paths are present, but in the case of ratio (A.1) one hundred twenty different paths are possible when changing five aggregates.[1] The proposed solution to resolve the ambiguity in the method of decomposition is to follow Fisher in creating his ideal index, and generalize the results by Gini (1937). In other words, one may take the geometric average of the different decompositions to obtain the different components.

The proposed decomposition of the yield index into five different index components, together with the efficiency index $EI = \frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \mathbf{y}_0)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \mathbf{y}_1)}$, obtains:

$$\frac{y_1^M/l_1}{y_0^M/l_0} = I^{(\mathbf{x}_1,\mathbf{x}_0)}(l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}; l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}) L^{(l_1,l_0)}(\mathbf{x}_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}; \mathbf{x}_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M})$$
$$Q^{(c_1,\mathbf{c}_0)}(\mathbf{x}_0, l_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}; \mathbf{x}_1, l_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}) K^{(\mathbf{b}_1,\mathbf{b}_0)}(\mathbf{x}_0, l_0, \mathbf{c}_0, \frac{\mathbf{y}_0^N}{y_0^M}; \mathbf{x}_1, l_1, \mathbf{c}_1, \frac{\mathbf{y}_1^N}{y_1^M})$$
$$Y^{(\frac{\mathbf{y}_1^N}{y_1^M}, \frac{\mathbf{y}_0^N}{y_0^M})}(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0; \mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1)\, EI, \tag{A.5}$$

where I is a variable-inputs component, L is a land-size component, Q is a soil-quality component, K is a capital-inputs component, and Y is an output-mix component.

Each index can be decomposed in different ways by varying the decomposition paths. Collecting the equal terms and following de Boer (2009) the index for variable

[1]The number of paths grows factorially with the number of aggregates.

inputs $\mathbf{x}$ is stated as follows:

$$I^{(\mathbf{x}_1,\mathbf{x}_0)}(l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M}) = \tag{A.6}$$

$$\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/5}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/5}.$$

This index shows the change in the maximum yields due to the change in agricultural inputs $\mathbf{x}$, while keeping the other inputs and outputs $(l_.,\mathbf{c}_.,\mathbf{b}_.,\frac{\mathbf{y}_.^N}{y_.^M})$ constant. In a similar manner one may obtain the land-size component:

$$L^{(l_1,l_0)}(\mathbf{x}_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M}) = \bar{L}^{(l_1,l_0)}(\mathbf{x}_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M})l_0/l_1 = \tag{A.7}$$

$$l_0/l_1\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/5}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/5}.$$

The soil-quality component is decomposed in the following way:

$$Q^{(c_1,\mathbf{c}_0)}(\mathbf{x}_0,l_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_1,l_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M})= \tag{A.8}$$

$$\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/5}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_1,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/30}\left\{\frac{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_1^N}{y_1^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_1,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_1,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_1,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}\left\{\frac{E(\mathbf{x}_0,l_1,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_1,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_0,l_0,\mathbf{c}_1,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}{E(\mathbf{x}_0,l_0,\mathbf{c}_0,\mathbf{b}_0,\frac{\mathbf{y}_0^N}{y_0^M},1)}\right\}^{1/5}.$$

The capital-inputs component is decomposed in the following manner:

$$K^{(\mathbf{b}_1,\mathbf{b}_0)}(\mathbf{x}_0,l_0,\mathbf{c}_0,\frac{\mathbf{y}_0^N}{y_0^M};\mathbf{x}_1,l_1,\mathbf{c}_1,\frac{\mathbf{y}_1^N}{y_1^M})= \tag{A.9}$$

$$\left\{\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}\right\}^{1/5} \left\{\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_1, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_0, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30} \left\{\frac{E(\mathbf{x}_1, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_1, l_0, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30} \left\{\frac{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}\right\}^{1/30} \left\{\frac{E(\mathbf{x}_0, l_1, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}\right\}^{1/30} \left\{\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_0, l_1, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/5}.$$

The output-mix component of the yield-index difference is decomposed as follows:

$$Y^{(\frac{\mathbf{y}_1^N}{y_1^M}, \frac{\mathbf{y}_0^N}{y_0^M})}(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0; \mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1) = \tag{A.10}$$

$$\left\{\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/5} \left\{\frac{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_1, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30}$$

$$\left\{\frac{E(\mathbf{x}_0, l_1, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30} \left\{\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30} \left\{\frac{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30}$$

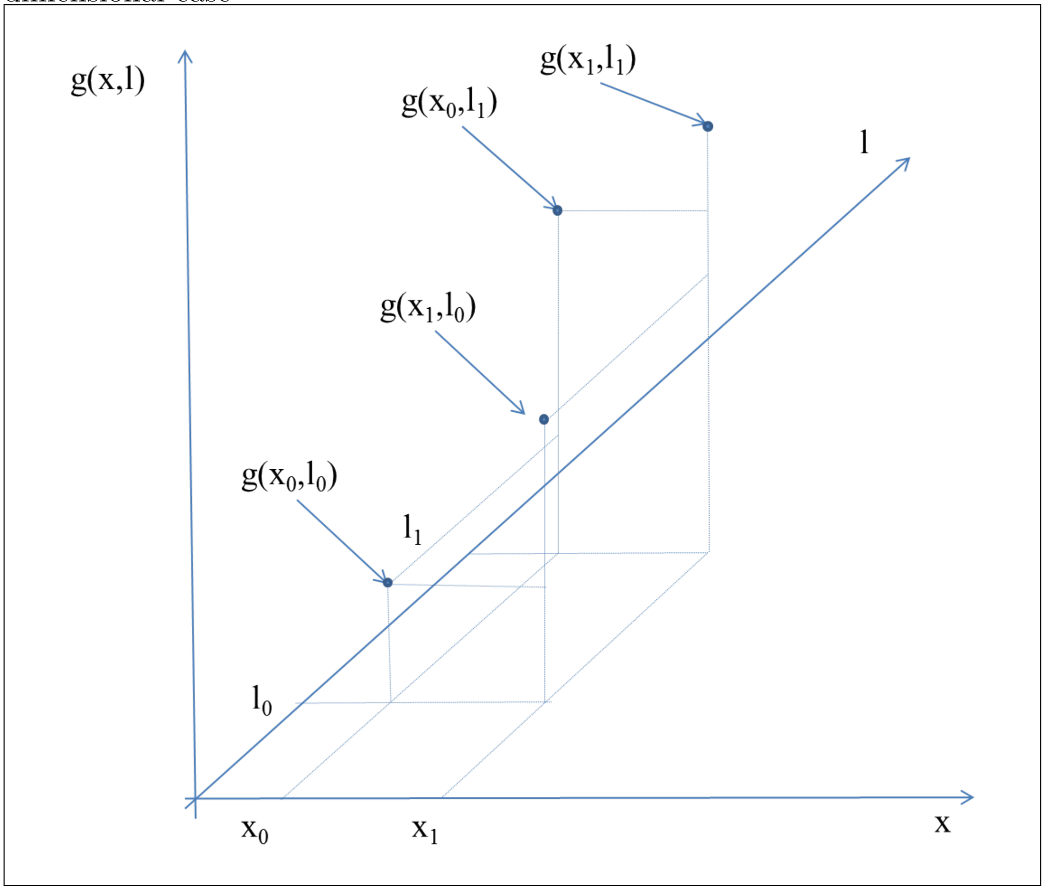
$$\left\{\frac{E(\mathbf{x}_1, l_0, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_0, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30} \left\{\frac{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/30} \left\{\frac{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_1, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20}$$

$$\left\{\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_1, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_1, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20} \left\{\frac{E(\mathbf{x}_0, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_1, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)}\right\}^{1/20}$$

$$\left\{ \frac{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_1^N}{y_1^M}, 1)}{E(\mathbf{x}_0, l_0, \mathbf{c}_0, \mathbf{b}_0, \frac{\mathbf{y}_0^N}{y_0^M}, 1)} \right\}^{1/5}.$$

A.1 Figures

Figure A.1: Graphical representation of the problem of path dependency in a two-dimensional case



Bibliography

A. Abdulai and C. Binder. Slash-and-burn cultivation practice and agricultural input demand and output supply. *Environment and Development Economics*, 11 (02):201–220, 2006. doi: 10.1017/S1355770X05002779.

J. Áczel. *Lectures in Functional Equations and Their Applications. Mathematics in Science and Engineering. Volume 19*. Elsevier, 1970.

D.A. Ali and K. Deininger. Is there a farm-size productivity relationship in African agriculture ? Evidence from Rwanda. Policy Research Working Paper Series 6770, The World Bank, February 2014.

J. Assunçao and L.H.B. Braido. Testing household-specific explanations for the inverse productivity relationship. *American Journal of Agricultural Economics*, 89(4):980–990, 2007.

R. Banker. Hypothesis tests using data envelopment analysis. *Journal of Productivity Analysis*, 7:139 – 160, 1996.

P.K. Bardhan. Size, productivity, and returns to scale: An analysis of farm-level data in Indian agriculture. *Journal of Political Economy*, 81(6):1370–86, Nov.-Dec. 1973.

C.B. Barrett. On price risk and the inverse farm size-productivity relationship. *Journal of Development Economics*, 51(2):193–215, 1996. PT: J; TC: 23.

C.B. Barrett, M.F. Bellemare, and J.Y. Hou. Reconsidering conventional explanations of the inverse productivity-size relationship. *World Development*, 38(1):88 – 97, 2010. ISSN 0305-750X. doi: DOI:10.1016/j.worlddev.2009.06.002.

M.R. Bellon and J.E. Taylor. 'Folk' Soil Taxonomy and the Partial Adoption of New Seed Varieties. *Economic Development and Cultural Change*, 41(4):pp. 763–786, 1993. ISSN 00130079.

D. Benjamin. Can unobserved land quality explain the inverse productivity relationship? *Journal of Development Economics*, 46(1):51 – 84, 1995. ISSN 0304-3878. doi: DOI:10.1016/0304-3878(94)00048-H.

E. R. Berndt and L. R. Christensen. The translog function and the substitution of equipment, structures, and labor in U.S. manufacturing 1929-68 . *Journal of Econometrics*, 1(1):81 – 113, 1973. ISSN 0304-4076.

A.R. Berry and W.R. Cline. *Agrarian structure and productivity in developing countries*. The Johns Hopkins University Press, 1979.

S.S. Bhalla. Does land quality matter? : Theory and measurement. *Journal of Development Economics*, 29(1):45 – 62, 1988. ISSN 0304-3878. doi: DOI:10.1016/0304-3878(88)90070-3.

S.S. Bhalla and P.L. Roy. Mis-specification in farm productivity analysis: The role of land quality. *Oxford Economic Papers*, 40(1):55–73, 1988. ISSN 00307653.

C. Blackorby, D. Primont, and R.R. Russell. On testing separability restrictions with flexible functional forms. *Journal of Econometrics*, 5(2):195 – 209, 1977. ISSN 0304-4076.

C. Blackorby, D. Primont, and R.R. Russell. *Duality, Separability and Functional Structure: Theory and Applications*. Amsterdam: North-Holland, 1978.

N.C. Brady and R.R. Weil. *The nature and properties of soils*. Pearson Prentice Hall, 2008.

C. Carletto, S. Savastano, and A. Zezza. Fact or artifact: The impact of measurement errors on the farm size-productivity relationship. *Journal of Development Economics*, 103(0):254 – 261, 2013. ISSN 0304-3878. doi: http://dx.doi.org/10.1016/j.jdeveco.2013.03.004.

M.R. Carter. Identification of the inverse relationship between farm size and productivity - an empirical-analysis of peasant agricultural production. *Oxford Economic Papers-New Series*, 36(1):131–145, 1984. PT: J; TC: 29.

D.W. Caves, L.R. Christensen, and W.E. Diewert. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 50(6):pp. 1393–1414, 1982. ISSN 00129682.

H.H. Chang and F. Wen. Off-farm work, technical efficiency, and rice production risk in Taiwan. *Agricultural Economics*, 42(2):269–278, 2011. ISSN 1574-0862. doi: 10.1111/j.1574-0862.2010.00513.x.

A.V. Chayanov. *The theory of peasant economy*. The University of Wisconsin Press published in 1986, 1926.

C. Daraio, L. Simar, and P.W. Wilson. Testing whether two-stage estimation is meaningful in non-parametric models of production. *Discussion Paper 1031, Institut de Statistique, UCL, Louvain-la-Neuve, Belgium*, 2010.

P. de Boer. Generalized Fisher index or Siegel Shapley decomposition? *Energy Economics*, 31(5):810 – 814, 2009. ISSN 0140-9883. doi: 10.1016/j.eneco.2009.03.003.

S. Di Falco and J.P. Chavas. On crop biodiversity, risk exposure, and food security in the Highlands of Ethiopia (2008-10). *American Journal of Agricultural Economics*, 91(3):599–611, 2009. doi: 10.1111/j.1467-8276.2009.01265.x.

W.E. Diewert and T.J. Wales. Flexible functional forms and tests of homogeneous separability. *Journal of Econometrics*, 67(2):259–302, June 1995.

S. Emran and I. Alam. Weak separability of non-tradables from consumer good imports: A simple test with evidence from Bangladesh. *Economics Letters*, 63(2): 225–234, 1999.

R. Färe, S. Grosskopf, M. Norris, and Z. Zhang. Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review*, 84(1):pp. 66–83, 1994. ISSN 00028282.

R. Färe, S. Grosskopf, and P. Roos. Productivity and quality changes in Swedish pharmacies. *International Journal of Production Economics*, 39(1-2):137–144, April 1995.

R. Färe, S. Grosskopf, and F. Hernandez-Sancho. Environmental performance: an index number approach. *Resource and Energy Economics*, 26(4):343 – 352, 2004. ISSN 0928-7655. doi: http://dx.doi.org/10.1016/j.reseneeco.2003.10.003.

G. Feder. The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints. *Journal of Development Economics*, 18(23):297 – 313, 1985. ISSN 0304-3878. doi: http://dx.doi.org/10.1016/0304-3878(85)90059-8.

C.E. Ferguson. *The Neoclassical Theory of Production and Distribution.* Number 9780521076296 in Cambridge Books. Cambridge University Press, 2008.

I. Fisher. *The Making of Index Numbers: A Study of Their Varieties, Tests, and Reliability.* Houghton Mifflin, 1922.

N. Fuwa, C. Edmonds, and P. Banik. Are small-scale rice farmers in eastern India really inefficient? Examining the effects of microtopography on technical efficiency estimates. *Agricultural Economics*, 36(3):335–346, 2007. ISSN 1574-0862. doi: 10.1111/j.1574-0862.2007.00211.x.

D.O. Gilligan. Farm size, productivity, and economic efficiency: accounting for differences in efficiency of farms by size in Honduras. *Unpublished paper presented at the 1998 American Agricultural Economics Association Annual Meetings*, pages 1 – 14, 1998.

C. Gini. Methods of eliminating the influence of several groups of factors. *Econometrica*, 5(1):pp. 56–73, 1937. ISSN 00129682.

A. Hailu and R.G. Chambers. A Luenberger soil-quality indicator. *Journal of Productivity Analysis*, 38:145–154, 2012. ISSN 0895-562X.

S.M. Helfand and E.S. Levine. Farm size and the determinants of productive efficiency in the Brazilian Center-West. *Agricultural Economics*, 31(2 - 3):241 – 249, 2004.

D.J. Henderson and R.R. Russell. Human capital and convergence: A production-frontier approach. *International Economic Review*, 46(4):1167–1205, 2005. ISSN 1468-2354. doi: 10.1111/j.1468-2354.2005.00364.x.

K.G. Husbands, T.O. Konyango, and T.C. Pinckney. Education and agricultural productivity in Kenya. *In "The evaluation of public expenditure in Africa", eds. H.J. Bruton, C. Hill, World Bank Publication, Economic Development Institute, Washington, D.C.*, pages 111–138, 1996.

E.C. Jaenicke and L.L. Lengnick. A soil-quality index and its relationship to efficiency and productivity growth measures: Two decompositions. *American Journal of Agricultural Economics*, 81(4):881–893, 1999. ISSN 00029092.

D.W. Jorgenson and Z. Griliches. The explanation of productivity change. *The Review of Economic Studies*, 34(3):249–283, 1967. ISSN 00346527.

A. Kimhi. Plot size and maize productivity in Zambia: is there an inverse relationship? *Agricultural Economics*, 35(1):1 – 9, 2006.

A. Kneip, L. Simar, and P. Wilson. Central Limit Theorems for DEA and FDH Efficiency Scores: When Bias Kills the Variance. *Econometric Theory*, 31(2): 394–422, 2015a.

A. Kneip, L. Simar, and P.W. Wilson. Testing hypotheses in non-parametric models of production. *Discussion Paper 48, Institut de statistique, biostatistique et sciences actuarielles (ISBA), UCL, Louvain-la-Neuve, Belgium, forthcoming in Journal of Business and Economic Statistics*, 2015b.

A. Kolmogorov. Confidence limits for an unknown distribution function. *The Annals of Mathematical Statistics*, 12(4):pp. 461–463, 1941. ISSN 00034851.

S. Kumar and R.R. Russell. Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence. *American Economic Review*, 92(3):527–548, JUN 2002. PT: J.

R.L. Lamb. Inverse productivity: land quality, labor markets, and measurement error. *Journal of Development Economics*, 71(1):71 – 95, 2003. ISSN 0304-3878. doi: DOI:10.1016/S0304-3878(02)00134-7.

A. Lewbel and O. Linton. Nonparametric matching and efficient estimators of homothetically separable functions. *Econometrica*, 75(4):1209–1227, 2007. ISSN 1468-0262.

Q. Li. Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews*, 15(3):261–274, 1996.

Q. Li, E. Maasoumi, and J.S. Racine. A nonparametric test for equality of distributions with mixed categorical and continuous data. *Journal of Econometrics*, 148 (2):186–200, February 2009.

P. Loveland and J. Webb. Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. *Soil and Tillage Research*, 70(1):1 – 18, 2003. ISSN 0167-1987.

P. Marenya and C.B. Barrett. Soil quality and fertilizer use rates among smallholder farmers in western Kenya. *Agricultural Economics*, 40(5):561–572, 2009a.

P.P. Marenya and C.B. Barrett. State-conditional fertilizer yield response on western Kenyan farms. *American Journal of Agricultural Economics*, 91(4):991–1006, 2009b.

O. B. Olesen. A homothetic reference technology in data envelopment analysis. *European Journal of Operational Research*, 233(3):759 – 771, 2014. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/j.ejor.2013.09.024.

F. Piette. Les déterminants de la productivité agricole dans le nord-est du Brésil. *Université de Montéral. Départment d'économie*, 2006.

D. Primont and D. Primont. Homothetic non-parametric production models. *Economics Letters*, 45(2):191 – 195, 1994. ISSN 0165-1765.

G.R. Saini. Holding size, productivity, and some related aspects of Indian agriculture. *Economic and Political Weekly*, 6(26):pp. A79+A81–A85, 1971. ISSN 00129976.

T. Schubert and L. Simar. Innovation and export activities in the german mechanical engineering sector: an application of testing restrictions in production analysis. *Journal of Productivity Analysis*, pages 1–15, 2010. ISSN 0895-562X. DOI: 10.1007/s11123-010-0199-6.

A. Sen. An Aspect of Indian Agriculture. *Economic Weekly*, 14:243–246, 1962.

S.M. Sherlund, C.B. Barrett, and A.A. Adesina. Smallholder technical efficiency controlling for environmental production conditions. *Journal of Development Economics*, 69(1):85 – 101, 2002. ISSN 0304-3878. doi: DOI:10.1016/S0304-3878(02) 00054-8.

L. Simar and P.W. Wilson. Non-parametric tests of returns to scale. *European Journal of Operational Research*, 139(1):115 – 132, 2002. ISSN 0377-2217. doi: DOI:10.1016/S0377-2217(01)00167-9.

L. Simar and P.W. Wilson. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1):31–64, Jan 2007.

L. Simar and P.W. Wilson. Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Productivity Analysis*, pages 1–21, 2010. ISSN 0895-562X. DOI: 10.1007/s11123-010-0200-4.

L. Simar and V. Zelenyuk. On testing equality of distributions of technical efficiency scores. *Econometric Reviews*, 25(4):497–522, 2006.

W.J. Spillman. The efficiency movement in its relation to agriculture. *Annals of the American Academy of Political and Social Science*, 59:pp. 65–76, 1915. ISSN 00027162.

M. Verschelde, M. D'Haese, E. Vandamme, and G. Rayp. Challenging small-scale farming, a non-parametric analysis of the (inverse) relationship between farm productivity and farm size in Burundi. *AgEcon Search*, pages 1–12, 2011.

P.W. Wilson. Detecting outliers in deterministic nonparametric frontier models with multiple outputs. *Journal of Business and Economic Statistics*, 11(3):319–323, 1993.

A.D. Woodland. On testing weak separability. *Journal of Econometrics*, 8(3):383–398, December 1978.

K.D. Zieschang. On the structure of technologies of an input-limited unit. *In W. Eichhorn, R. Henn, K. Neumann, and R.W. Shepard, Eds.: Quantitative studies on production and prices, Wurzburg: Physica Verlag*, pages 57 – 69, 1983.