Environment for Development

Discussion Paper Series June 2017 ■ EfD DP 17-06

Productive Efficiency and Its Determinants in a Changing Climate

A Monotonic Translog Stochastic Frontier Analysis

Richard Mulwa and Jane Kabubo-Mariara

Environment for Development Centers

Central America
Research Program in Economics and Environment for Development in Central America Tropical Agricultural Research and Higher Education Center (CATIE)

Chile
Research Nucleus on Environmental and Natural Resource Economics (NENRE)
Universidad de Concepción

China
Environmental Economics Program in China (EEPC)
Peking University

Colombia
The Research Group on Environmental, Natural Resource and Applied Economics Studies (REES-CEDE), Universidad de los Andes, Colombia

Ethiopia
Environment and Climate Research Center (ECRC)
Ethiopian Development Research Institute (EDRI)

India
Centre for Research on the Economics of Climate, Food, Energy, and Environment, (CECFEE), at Indian Statistical Institute, New Delhi, India

Kenya
School of Economics
University of Nairobi

South Africa
Environmental Economics Policy Research Unit (EPRU)
University of Cape Town

Sweden
Environmental Economics Unit
University of Gothenburg

Tanzania
Environment for Development Tanzania
University of Dar es Salaam

USA (Washington, DC)
Resources for the Future (RFF)

Vietnam
University of Economics
Ho Chi Minh City, Vietnam

The Environment for Development (EfD) initiative is an environmental economics program focused on international research collaboration, policy advice, and academic training. Financial support is provided by the Swedish International Development Cooperation Agency (Sida). Learn more at www.efdinitiative.org or contact info@efdinitiative.org.

Productive Efficiency and Its Determinants in a Changing Climate: A Monotonic Translog Stochastic Frontier Analysis

Richard Mulwa and Jane Kabubo-Mariara

Abstract

The changing weather patterns and seasonal shifts are negatively impacting agricultural ecosystems and compromising the benefits from production of agricultural goods and services. Such impacts include reduced farm returns, reduced household incomes, increase in poverty levels, and reduction in farm productivity and efficiency. Using three waves of panel data, this study applies a monotonic translog stochastic frontier (SFA) to assess the overall farm efficiency and the influence of climatic factors, agro-ecological factors, and household factors on farm level efficiency. From the results, farming households are, on average, 63% efficient and could expand output by about 37% and still use the same level of inputs. However, this will be determined by a number of exogenous determinants such as climatic, agro-ecological, and household factors. Climatic factors such as rainfall and temperature decrease and increase inefficiency, respectively. The quasi-fixed factors of education and household size decrease and increase inefficiency, respectively, while age of the household head decreases inefficiency.

Key Words: climate change and variability, farm productive efficiency, monotonic translog stochastic frontier

JEL Codes: Q18, Q54

Contents

Productive Efficiency and Its Determinants in a Changing Climate: A Monotonic Translog Stochastic Frontier Analysis

Richard Mulwa and Jane Kabubo-Mariara*

1. Introduction

In Kenya, climate change is causing variability in weather factors such as precipitation, temperature and soil moisture, and some areas in the country have experienced prolonged droughts and increased temperatures, which have compromised production of food crops (Mulwa et al. 2016). The projected future impacts could include reduced farm productivity and efficiency, reduced farm returns, reduced household incomes, and an increase in poverty levels (Mulwa et al. 2016). In a conventional rainfed-production system, agricultural households use direct factors of production (fertilizer, seed, labour, etc.) to produce several agricultural outputs. However, in a changing climate, production is also influenced by climate factors (precipitation, temperature, soil moisture, etc.). Other factors that influence production include household characteristics and agro-ecological characteristics. Climatic factors and agro-ecological factors are outside the control of farming households, while production inputs and most household characteristics are within the control of the household. The farmer has to decide how to combine production inputs subject to his/her exogenous environment. The outcome of the farmer's decisions will influence how much will be produced, and hence production efficiency.

Against this backdrop, this study seeks to answer two key questions: (a) how efficient are agricultural households in Kenya in a changing climate setting; and (b) how do household factors, agro-ecological factors, and climate factors influence farm production. The study also asks what policy decisions can be drawn from these results to aid in improving efficiency in a changing climate.

Despite the increasing number of climate change and variability studies and productive efficiency studies, there is a dearth of literature linking climate change and variability to farm-level productive efficiency in Africa. This study makes a significant

* Richard Mulwa, Ph.D., CASELAP, University of Nairobi, P.O. Box 30197-00100, Nairobi. Email: richard.mulwa@uonbi.ac.ke. Jane Kabubo-Mariara, Ph.D., School of Economics, University of Nairobi, P.O. Box 30197-00100, Nairobi. Email: jmariara@uonbi.ac.ke.

contribution to this limited literature by linking the two fields. This is done by incorporating climatic factors, namely precipitation, temperature, and the Standardized Precipitation and Evapo-transpiration index (SPEI); household characteristics, i.e., household size, education of household head, etc.; and altitude, which is a proxy for agro-ecological zones. Productive efficiency is estimated by use of a monotonic translog stochastic frontier (SFA) model.

Since the publication of Farrell's article on measurement of productive efficiency in 1957, a number of techniques – both parametric and non-parametric – have been used to estimate efficiency of production. The non-parametric methods include Data Envelopment Analysis (Emrouznejad and Podnovski 2004; Mulwa et al. 2009) and Free Disposable Hull (Thrall 1999; Cherchye et al. 2000) and use mathematical programming techniques to evaluate the performance of firms or decision making units (DMUs). Parametric methods are either stochastic or deterministic and use econometric techniques for estimation (Kumbakhar and Lovell 2000; Coelli et al. 2005).

The main stochastic frontier methods include the Cobb-Douglas and translog stochastic frontier models (O'Donnell and Coelli 2005; Sauer et al. 2006). The main limitation of the Cobb-Douglas frontier is that it uses only a few parameters and is simplistic in that it assumes all firms have same production elasticities and that substitution elasticities are equal to one (Battese and Coelli 1992). The translog stochastic frontier model has advantages over the Cobb-Douglas because it's more flexible and deals with most setbacks presented by the Cobb-Douglas (Coelli et al. 2005). The introduction of the monotonicity in the translog model handles the rare cases of a negative technical input–output relationship because it is always assumed that production functions monotonically increase in all inputs, i.e., the output quantity must not decrease if any input quantity is increased (Henningsen and Henning 2009). This justifies the choice of the monotonic translog stochastic frontier model over its counterparts.

Overall, the approaches used in this study are advantageous over most existing methodologies in similar studies for a variety of reasons. First, the study uses three waves of panel data, which allows us to exploit the time dimension, unlike cross-sectional data, which considers one period. Second, the use of the monotonic translog stochastic frontier is methodologically innovative, as the model has not been used in most translog stochastic frontier analysis. Further, the farm output is measured in kilocalories produced on the farm, unlike most studies, which estimate efficiency using a single output and multiple inputs. Finally, our study is country-wide and considers all cropping activities on the farm and is therefore replicable elsewhere in the country and also in other developing countries.

2. Literature Review

Agriculture is affected in complex ways by climate change and variability (IPCC 2007). The number of studies on the impact of climate change and variability on crop and livestock productivity has increased over the decades, with a more recent focus on developing countries in general, and specifically Africa. Most of the studies assess the extent to which adaptation options can lessen the expected impact of climate change and variability. The bulk of the emerging studies have used the Ricardian model, following Mendelsohn et al. (1994). By regressing land values (or net revenue) on climate normals, e.g., 30 year averages of climate variables rather than weather factors and other exogenous factors, one can then determine the marginal contribution of each input to farm income as capitalized in land value (see, for instance, Gbetibouo and Hassan 2005; Deressa et al. 2009; Kabara and Kabubo-Mariara 2011). Other studies use similar approaches to the Ricardian method, but investigate the impact of climate change on nutrition (see, for instance, Kabubo-Mariara and Kabara 2015; Kabubo-Mariara et al. 2016; 2017).

Despite the increasing number of climate change and variability studies, there is a dearth of literature linking weather and climate change factors and farm-level productive efficiency in Africa. Some of the studies have been inspired by Farrell (1957), who asserted that environmental factors such as air and water quality, climate and location need attention because they affect farm productive efficiency. In Kenya, using a two-stage semi-parametric approach on household panel data, Ogada et al. (2014) reveal that technical efficiency differentials in smallholder farms are influenced by environmental factors, production risks, and farmer characteristics. In addition, Mulwa and Kabubo-Mariara (2017) use a two-stage semi-parametric approach to assess the influence of climatic, household, and institutional factors on farm efficiency in Kenya. The study revealed that these factors influence productive efficiency either positively or negatively. The policy implication from the two studies is that Kenya has room to improve agricultural productivity by addressing climatic, adaptation and farm-level constraints. Oyekale (2012), using a stochastic frontier model, found that climate change poses serious problems to Nigerian cocoa production. The presence of a high level of inefficiency in smallholder farm production in Sub-Saharan Africa has also been attributed to a number of other factors, such as limited use of inputs, limited market access and household characteristics (Mkhabela et al. 2010; Alemu et al. 2009).

A study on Ethiopian smallholder farms showed that seasonal climate conditions (including rainfall and temperature) and agro-ecological setting have significant impact on technical efficiency in Ethiopian agriculture (Alemu et al. 2009). The study also observed that education, proximity to markets, and access to credit contribute to significant reduction in farm inefficiency. In this regard, it is necessary to understand the influence of socio-economic characteristics, management practices and environmental factors on farm productive efficiency.

Outside Africa, a number of studies explore the effect of precipitation on food production and farm efficiency. Makki et al. (2012) evaluated the impact of climate change on productivity and technical efficiency of paddy farms in tidal swamp-land in Indonesia. The analysis showed that climate change has a negative effect on farm efficiency. Brazdik (2006) groups factors influencing farm efficiency into three broad categories: farm-specific variables (intensity of inputs such as labour, fertilizers and seeds, farm size, organizational structure such as tenure, and crop variety); economic factors (inputs prices); and environmental factors (rainfall, temperature or wet–dry periods, and agro-ecological zone). Controlling for environmental factors in technical efficiency analysis of smallholder production across different ecological zones has been shown to improve the precision of results (Sherlund et al. 2002). The strength of the agro-ecological argument is reinforced by studies analyzing district-level crop yield and precipitation in India (Asada and Matsumoto 2009; Kumar et al. 2004).

Lobell and Burke (2008) report that a change in growing season precipitation can be associated with a significant change in production efficiency. This is supported by Henderson and Kingwell (2005), who show that rainfall, along with purchased inputs, such as labour and materials, are important inputs to wool production in Australia. In India, large decreases in efficiency of crop productivity have been attributed to anomalously low precipitation events and extremely high temperatures (Kumar et al. 2004).

Despite the increasing number of climate change studies, there is a dearth of literature linking climate change and variability factors and farm-level productive efficiency in developing countries, where climate change is expected to have more adverse impact. In addition, the influence of climatic and agro-ecological factors on farm productive efficiency remains less explored in the case of Kenya. Save for Ogada et al.

(2014), studies on Kenya have generally examined the influence of farm-specific and economic factors on agricultural efficiency (see Kirimi and Swinton 2004; Kibaara 2005; Nyagaka et al. 2010; Mulwa et al. 2009). This study is therefore an addition to the limited literature on climate change and efficiency. The study has additional advantages over existing studies on climate change impacts in that it introduces climate variability (weather) and conventional factors of production, which are not considered in a classical Ricardian analysis. The study also introduces a methodological innovation by using a translog stochastic frontier production function with imposed monotonicity, which has not been considered in most stochastic frontier studies.

3. Methodology

3.1. Theoretical and Empirical Approaches

The stochastic production frontier was first proposed independently by Aigner et al. (1977) and Meeusen and van den Broeck (1977). Their general setup is

$$y_i = m(\boldsymbol{x_i}, \beta) - u_i + v_i \implies m(\boldsymbol{x_i}, \beta) - \varepsilon_i \tag{1}$$

where the key difference from a standard production function is the appearance of the two distinct error terms in the model. The u_i term captures inefficiency, defined as shortfall from maximal output dictated by the production function, $m(\boldsymbol{x_i}, \beta)$, while the v_i terms capture outside influences beyond the control of the producer.

To estimate Equation (1) via maximum likelihood, the density of ε must be determined. Once distributional assumptions on v and u have been made, $f(\varepsilon)$ can be determined by noting that the joint density of u and v, $f(u, v)$, can be written as the product of the individual densities, $f(u)f(v)$, given the independence of u and v.

Further, because $v = \varepsilon + u$, then $f(u, \varepsilon) = f(u)f(\varepsilon + u)$ and u should be integrated out to obtain $f(\varepsilon)$. However, note that not all distributional assumptions will provide closed form solutions for $f(\varepsilon)$. With either the half-normal specification of Aigner et al. (1977) or the exponential specification of Meeusen and van den Broeck (1977), $f(\varepsilon)$ possesses an (approximately) closed form solution, making direct application of maximum likelihood straightforward. The density of ε for the half-normal is given as

$$f(\varepsilon) = \frac{2}{\sigma}\phi(\varepsilon/\sigma)\Phi(-\varepsilon\lambda/\sigma) \tag{2a}$$

while that of the normal exponential is given by

$$f(\varepsilon) = \frac{1}{\sigma_u} \Phi(-\varepsilon/\sigma_v - \sigma_v/\sigma_u) e^{\varepsilon/\sigma_u + \sigma_v^2/2\sigma_u^2} \quad (2b)$$

where $\phi(.)$ is the standard normal probability density function, $\Phi(.)$ is the standard normal cumulative distribution function, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$ and $\lambda = \sigma_u/\sigma_v$. The parameterization is quite common and has intuitive appeal, and λ can be thought of as a measure of the inefficiency-to-noise ratio, i.e., the amount of variation in ε due to inefficiency versus the variation due to noise. Based on Battese and Corra (1977), parameterization $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$. The parameter, γ, must lie between 0 and 1 and can be used as a quick check of the presence of inefficiency.

The half-normal assumption for the one-sided inefficiency term is undoubtedly the most commonly used in empirical studies of inefficiency. However, a variety of alternative stochastic frontier models have been proposed, using alternative distributions on the one-sided term. Most notably, Stevenson (1980) proposed a generalization of the half-normal distribution, the truncated (at 0) normal distribution. The truncated normal distribution depends on two parameters (μ, σ_u^2) and affords the researcher more flexibility in the shape of the distribution of inefficiency. The truncated normal is given as

$$f(u) = \frac{1}{\sqrt{2\pi}\sigma_u \Phi(\mu/\sigma_u)} e^{-\frac{(u-\mu)^2}{2\sigma_u^2}} \quad (3)$$

When $\mu = 0$, this distribution reduces to the half-normal. Unlike the half-normal and exponential densities, the truncated normal density has its mode at 0 only when $\mu \leq 0$, but otherwise has its mode at μ. Thus, for $\mu > 0$, the implication is that, in general, the DMU is inefficient. The density of ε for the normal truncated normal specification is

$$f(\varepsilon) = \frac{1}{\sigma} \phi\left(\frac{\varepsilon+\mu}{\sigma}\right) \Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\varepsilon\lambda}{\sigma}\right) / \Phi(\mu/\sigma_u) \quad (4)$$

In addition, Greene (1980; 1990) and Stevenson (1980) both proposed and used a gamma distribution for inefficiency.

Assuming a single-output, multiple-input cross-sectional setup, we can assume the translog production function, which is conceptually simple and imposes no a priori restrictions on the structure of the technology. It also satisfies second-order flexibility (Diewert 1974) and its logarithmic form has the advantage that inefficiencies are captured by an additive rather than a multiplicative term (Henningsen and Henning 2009). The stochastic frontier form of the function (Greene 1980; Kaliranjan 1990) can be stated as

$$lny = lnf(\boldsymbol{x}, \boldsymbol{\beta}) = \beta_0 + \sum_{i=1}^{N} \beta_i \, lnx_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \beta_{ij} \, lnx_i lnx_j + v_i - u_i \quad (5)$$

where $i = 1,2, \ldots, N$ are cross-sectional units (or decision-making units); $i, j = 1,2, \ldots, N$ are the production inputs; lny is the natural log of the output of the m^{th} DMU; and lnx_i is the natural log of the input i of the m^{th} DMU. From Young's Theorem, $\beta_{ij} = \beta_{ji}$ and the marginal products are given as

$$f_i = \frac{f(\boldsymbol{x},\boldsymbol{\beta})}{x_i}\left(\beta_i + \sum_{j=1}^{N} \beta_{ij} lnx_j\right) \tag{6}$$

and the second-order derivative is given by

$$f_i = \frac{f(\boldsymbol{x},\boldsymbol{\beta})}{x_i x_j}\left((\beta_i + \sum_{k=1}^{N} \beta_i lnx_k).\left(\beta_j + \sum_{k=1}^{N} \beta_{jk} lnx_k - \Delta_{ij}\right) - \beta_{ij}\right) \tag{7}$$

where Δ_{ij} is the Kronecker delta, with $\Delta_{ij} = 1$ if $i = j$ and $\Delta_{ij} = 0$ otherwise (Henningsen and Henning 2009). Because all input quantities are non-negative and the translog functional form guarantees that output is positive, the monotonicity conditions of the translog function reduce to

$$\beta_i + \sum_{j=1}^{N} \beta_{ij} lnx_j \geq 0 \quad \forall \ i \tag{8}$$

The rationale for monotonicity from economic theory is that, in rare cases of negative input-output relationship (e.g., excessive fertilizer harming the crops), it would be wise to refrain from using as much fertilizer, i.e., increasing the (unused) quantity of the input would at least leave the output unchanged (Henningsen and Henning 2009). Therefore, if a production frontier is not monotonically increasing, the efficiency estimates of the individual decision making units (DMUs) cannot be reasonably interpreted. Non-monotonicity would also distort the estimation of the exogenous determinant of efficiency, i.e., the technical efficiency effects model, as proposed by Batesse and Coelli (1995). To impose monotonicity, Henningsen and Henning (2009) suggest a three-step procedure based on Koebel et al. (2003)'s two-step procedure. The three steps involve: i) estimating the unrestricted stochastic frontier; ii) obtaining the restricted $\boldsymbol{\beta}$ parameters by minimum distance estimation; and iii) determining the efficiency estimates of the DMUs and the effects of exogenous determinants. This procedure is used in the current study.

3.2. Estimating Efficiency

After the model parameters are estimated, we can proceed to estimate observation-specific efficiency, which is one of the main interests of a stochastic frontier model. The estimated efficiency levels can be used to rank producers, identify under-performing producers, and determine which firms use best practices. This information is useful in helping to design public policy or subsidy programs aimed at improving the overall efficiency level of decision making units. Thus far, we have information on how to estimate σ_u^2, which is the unconditional mean of u_i i.e., it provides information regarding the shape of the half-normal distribution on u_i. This information would suffice if the interest were in the average level of technical inefficiency in the sample. However, if interest lies in the level of inefficiency of a given decision making unit (DMU), knowledge of σ_u^2 is not enough because it does not contain any individual-specific information. To overcome this problem, Jondrow, Lovell, Materov and Schmidt (JLMS) (1982) proposed an estimate u_i from the expected value of u_i conditional on the composed error of the model ($\varepsilon_i \equiv v_i - u_i$). This conditional mean of u_i given ε_i gives a point estimate of u_i. The composed error contains individual-specific information, and so the conditional expectation yields the observation-specific value of the inefficiency. This is like extracting signal from noise.

JLMS (1982) show that the conditional density function of u_i given $\varepsilon_i, f(u_i|\varepsilon_i)$ is $N_+(\mu_{*i}, \sigma_*^2)$, where

$$\mu_{*i} = \frac{-\varepsilon_i \sigma_u^2}{\sigma^2} \tag{9a}$$

$$\sigma_*^2 = \frac{\sigma_v^2 \sigma_u^2}{\sigma^2} \tag{9b}$$

From these, the conditional mean can be estimated as

$$E(u_i|\varepsilon_i) = \mu_{*i} + \frac{\sigma.\phi(\frac{\mu_{*i}}{\sigma_*})}{\Phi(\frac{\mu_{*i}}{\sigma_*})} \tag{9c}$$

Maximum likelihood estimates of the parameters are substituted into the equation to obtain estimates of firm-level inefficiency. This estimator will produce values that are guaranteed to be non-negative. Because the conditional distribution of u is known, one can derive moments of any continuous function of $u|\varepsilon$. The challenge with the JLMS efficiency estimator is that it is inconsistent because $n \longrightarrow \infty$. This is because, in a cross-section, as $n \longrightarrow \infty$ we have new firms being added to the sample with their own level of inefficiency instead of new observations to help determine a given firm's specific level of inefficiency. Second, the JLMS efficiency estimator is not designed to estimate unconditional inefficiency; it is designed to estimate inefficiency conditional on ε, for which it is a consistent estimator. Finally, the JLMS inefficiency estimator is
is a shrinkage estimator, and, on average, we overstate the inefficiency level of a firm with small u_i , while we understate inefficiency for a firm with large u_i.

The same technique as used in JLMS (1982) can be used to obtain observation-specific estimates of the efficiency index, e^{-u_i}. Battese and Coelli (1988) show that

$$E(e^{-u_i}|\varepsilon_i) = e^{-(\mu_{*i}+\frac{1}{2}\sigma_*^2)} \frac{\Phi\left(\frac{\mu_{*i}}{\sigma_*}-\sigma_*\right)}{\Phi\left(\frac{\mu_{*i}}{\sigma_*}\right)} \tag{10}$$

This estimator is bounded between 0 and 1, with a value of 1 indicating a fully efficient firm. Similar expressions for the JLMS (1982) and Battese and Coelli (1988) efficiency scores can be derived under the assumption that u is exponential, truncated normal and Gamma (Kumbhakar and Lovell 2000).

3.3. Stochastic Frontier and Panel Data

Schmidt and Sickles (1984) mention three problems with cross-sectional models that are used to measure inefficiency. First, the ML method, used to estimate parameters and the inefficiency estimates using the JLMS formula, depends on distributional assumptions for the noise and the inefficiency components. Second, the technical inefficiency component has to be independent of the regressor(s) – an assumption that is unlikely to be true if firms maximize profit and inefficiency is known to the firm (Mundlak 1961). Third, the JLMS estimator is not consistent, in the sense that the conditional mean or mode of $(u|\varepsilon)$ never approaches u as the number of firms (cross-sectional units) approaches infinity. However, if we have panel data, some of these rigidities can be removed. However, to overcome some of these limitations, the panel models make other assumptions, some of which may or may not be realistic.

If we consider a *time-invariant technical inefficiency* model – a case in which inefficiency is assumed to be constant over time and individual-specific – the unobservable individual effects of the classic panel data model is the base from which inefficiency is measured. The model can be expressed as

$$y_{it} = m(\boldsymbol{x}_{it}, \beta) + \varepsilon_{it} \tag{11}$$

$$\varepsilon_{it} = v_{it} - u_i \qquad u_i \geq 0,$$
$$i = 1, \dots, N \quad t = 1, \dots, T$$

where $m(\boldsymbol{x}_{it}, \beta)$ is a linear in parameters function of the variables in the vector $\boldsymbol{x}_{it}$, and $u_i \geq 0$ is the time-invariant technical inefficiency of individual i. This model utilizes the panel feature of the data via u_i which is specific to an individual and does not change over time. The stochastic panel model with time-invariant inefficiency can be estimated under either the fixed effects or random effects framework (Wooldridge 2010). Which framework to select depends on the level of relationship one is willing to assume between inefficiency and the covariates of the model. Under the fixed effects framework, correlation is allowed between $\boldsymbol{x}_{it}$ and u_i, whereas under the random effects framework, no correlation is present between $\boldsymbol{x}_{it}$ and u_i. Neither of these approaches requires distributional assumptions on u_i and are, thus, labeled as distribution-free approaches. In the fixed effects framework, we assume $f(.)$ is linear in $\boldsymbol{x}_{it}$ (e.g., the log of input quantities in a translog production function model). Therefore, the time-invariant inefficiency panel data stochastic frontier panel model can then be written as

$$\begin{aligned} y_{it} &= \beta_0 + \boldsymbol{x}'_{it}\boldsymbol{\beta} + v_{it} - u_i \\ &= (\beta_0 - u_i) + \boldsymbol{x}'_{it}\boldsymbol{\beta} + v_{it} \\ &= (\alpha_i) + \boldsymbol{x}'_{it}\boldsymbol{\beta} + v_{it} \end{aligned} \tag{12}$$

where $\alpha_i = \beta_0 - u_i$. Under the fixed effects framework, u_i and thus α_i are allowed to have arbitrary correlation with $\boldsymbol{x}_{it}$. The fixed effects panel methods yield consistent estimates of $\boldsymbol{\beta}$, but $\hat{\alpha}_i$ is a biased estimator of u_i because $u_i \geq 0$ by construction. Nevertheless, after $\hat{\alpha}_i$ is obtained, a simple transformation can be applied to recover $\hat{u}_i \geq 0$, which will be consistent provided $T \longrightarrow \infty$. A disadvantage of the FE approach is that no other time-invariant variables, such as gender, region, soils, etc., can be included in $\boldsymbol{x}_{it}$ because doing so entails perfect multicollinearity between α_i and the time-invariant regressors.

Under the Random Effects (RE) framework, it might be plausible to assume that α_i is uncorrelated with $\boldsymbol{x}_{it}$. When the assumption of no correlation between the covariates and firm inefficiency is indeed correct, then estimation of the stochastic frontier panel data model under the random effects framework provides more efficient estimates than estimation under the fixed effects framework. An important empirical advantage of the random effects framework is that time-invariant variables, such as gender, race, etc., may be included in the $\boldsymbol{x}_{it}$ vector of explanatory variables without leading to collinearity with α_i.

Under the *time-varying technical inefficiency models*, the implication is that an inefficient DMU learns over time. Models in which the inefficiency effects are time-varying are more general than the time-invariant models, in the sense that the time-invariant models can be viewed as special cases of the time-varying models. Battese and Coelli (1992) propose a stochastic frontier production function for (unbalanced) panel data which has firm effects which are assumed to be distributed as truncated normal random variables, which are also permitted to vary systematically with time. The model may be expressed as:

$$y_{it} = \beta_0 + \boldsymbol{x}'_{it}\boldsymbol{\beta} + v_{it} - u_{it}\,, \qquad i = 1, \dots, N,\ \ t = 1, \dots, T \tag{13}$$

where y_{it} is (the logarithm of) the production of the i^{th} DMU in the time period t; $\boldsymbol{x}_{it}$ is a vector of input quantities of the i^{th} DMU in the time period t; $\boldsymbol{\beta}$ is a vector of parameters to be estimated; v_{it} are random variables which are assumed to be $i.i.d \sim N(0, \sigma_v^2)$, and independent of the $u_{it} = u_i e^{(-\eta(t-T))}$, where the u_i are non-negative random variables which are assumed to account for technical inefficiency in production and are assumed to be truncations at zero of the $i.i.d \sim N(0, \sigma_u^2)$ distribution; and η is a parameter to be estimated. The panel of data need not be balanced. The imposition of one or more restrictions upon this model formulation can provide a number of the special cases of this particular model which have appeared in the literature. For instance, setting η to be zero provides the time-invariant model set out in Battese and Coelli (1988). The additional restriction of $\mu = 0$ reduces the model to the Pitt and Lee (1981) model. An additional restriction of $T = 1$ turns it into the original cross-sectional, half-normal formulation of Aigner et al. (1977).

Depending on any particular application, a large number of model choices could be considered. For example, does one assume a half-normal distribution or the more general truncated normal distribution for the inefficiency effects? If panel data is available (as is our case), should one assume time-invariant or time-varying efficiencies? When such decisions must be made, it is recommended that a number of the alternative models be estimated and that a preferred model be selected using likelihood ratio tests. In addition, one can test whether any form of stochastic frontier production function is required at all by testing the significance of the γ parameter. If the null hypothesis, that γ equals zero, is accepted, this would indicate that σ_u^2 is zero and the u_{it} term should be removed from the model, leaving a specification with parameters that can be consistently estimated using ordinary least squares.

3.4. Modelling Exogenous Determinants of Inefficiency

A question that is often asked is whether a DMU's level of inefficiency is dependent upon observable characteristics and, if so, how should this relationship be modeled in the context of a stochastic frontier? At various points in time, researchers have deployed a simpler, two-step analysis to model the influence of specific covariates on firm-level inefficiency. Pitt and Lee (1981) were the first to implement this type of approach, albeit in a slightly different form. Ali and Flinn (1989), Kalirajan (1990), Bravo-Ureta and Rieger (1991), Wollni and Brümmer (2012) and many others followed this two-step approach. This approach constructs estimates of observation-specific inefficiency via the Jondrow et al. (1982) conditional mean in the first step, and then regresses these inefficiency estimates on a vector of exogenous variables $\boldsymbol{z}_i$ in a second step. A negative coefficient of the exogenous variable in the regression is taken as an indication that firms with larger values of the variables tend to have a lower level of inefficiency (i.e., they are more efficient).

Criticisms against this two-step procedure have long pointed toward the biases that arise at various stages of the process; most prominently, the first stage model is misspecified (Battese and Coelli 1995). As explained in Wang and Schmidt (2002), if $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ are correlated, then the first step of the two-step procedure suffers from omitted variable bias. Even when $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ are uncorrelated, ignoring the dependence of the inefficiency on $\boldsymbol{z}_i$ will cause the estimated first-step technical efficiency index to be underdispersed, so that the results of the second-step regression are likely to be biased downward. In conclusion, this two-step analysis method has no statistical merit and duplication of this approach should be avoided.

The statistically acceptable one-step estimation method of investigating exogenous effects on inefficiency was first introduced in the truncated-normal model by Kumbhakar et al. (1991) and Reifschneider and Stevenson (1991). It was later deployed by Huang and Liu (1994) and Battese and Coelli (1995), each using slightly different algebraic forms for the pre-truncated parameterization of the mean function of u_i. All these studies assume that the mean of the distribution of the pre-truncated u_i is a linear function of the exogenous variables under investigation. That is, they abandon the constant-mean assumption on μ, and assume, instead, that the mean is a linear function of some exogenous variables, that is,

$$\mu_i = \boldsymbol{z}'_{\boldsymbol{ui}}\boldsymbol{\rho}^{\boldsymbol{u}} \qquad (14)$$

d $\boldsymbol{\rho}^{\boldsymbol{u}}$ is the corresponding coefficient vector As before, maximum likelihood estimation can be carried out to obtain estimates of $\boldsymbol{\rho}u$ along with all other model parameters. The Battese and Coelli (1995) model is equivalent to the Kumbhakar et al. (1991) specification, with the exceptions that allocative efficiency is imposed, the first-order profit maximising conditions are removed, and panel data is permitted. The Battese and Coelli (1995) model specification which is used in this analysis may be expressed as

$$y_{it} = \beta_0 + \boldsymbol{x}'_{it}\boldsymbol{\beta} + v_{it} - u_{it}\,, \qquad i = 1, \dots, N,\; t = 1, \dots, T \qquad (15)$$

where y_{it} is the log of the production of the i^{th} DMU in the time period t; $\boldsymbol{x}_{it}$ is a vector of input quantities of the i^{th} DMU in the time period t; $\boldsymbol{\beta}$ is a vector of parameters to be estimated; v_{it} are random variables which are assumed to be $i.i.d \sim N(0, \sigma_v^2)$; u_{it} are non-negative random variables which are assumed to account for technical inefficiency in production and are assumed to be $i.i.d \sim N(\mu_{it}, \sigma_u^2)$ truncations at zero, with:

$$\mu_{it} = \boldsymbol{z}'_{it}\boldsymbol{\rho} \tag{16}$$

where $\boldsymbol{z}_{it}$ is a vector of variables which may influence the efficiency of a firm; and $\boldsymbol{\rho}$ is a vector of parameters to be estimated.

4. Data

4.1. Description

This study used rural household survey data sets for the periods 2004, 2007, and 2010. Tegemeo Institute of Agricultural Policy and Development, Egerton University conducted the surveys in collaboration with Michigan State University.[1] The collected data were from 2,297, 1342 and 1313 households for the three years respectively, spread over 24 districts in Kenya. For our analysis, however, we used 3933 (1311 households per panel) from 22 districts across the country. The districts were drawn from six out of the eight provinces in the country. The selected districts cover four main agro-ecological zones, with sub-categorizations. These include coastal lowlands, lowlands, lower and upper midlands, and highlands (lower and upper). The districts vary in a range of agro-climatic conditions (i.e., rainfall, temperature, drought conditions/precipitation-evapotranspiration index, and elevation). The two districts which were dropped are located in the arid agro-ecological zones where crop production is hardly practiced (unless under irrigation) and the dominant activity is cattle ranching. The household data captures socio-economic characteristics including age, education, and household size. Household income sources, besides crop and livestock, include salary earning and individual business activities.

The survey data was also complemented by data on climate variables, namely, the SPEI (Standardized Precipitation-Evapotranspiration Index), rainfall and temperature. The SPEI is a multi-scaler drought index based on climatic data, which takes into account both precipitation and potential evapotranspiration in determining drought. It captures the main impact of increased temperatures on water demand. The SPEI data was obtained from the Global SPEI database and covers the period between January 1901 and

[1] http://www.tegemeo.org/.

December 2013 (Begueria and Vicente-Serrano 2013). Rainfall data was obtained from the CHIRPS (Climate Hazards Group InfraRed Precipitation with Stations) data archive. CHIRPS is a 30+ year quasi-global rainfall dataset, which incorporates satellite imagery with in-situ station data to create gridded rainfall time series for trend analysis and seasonal drought monitoring (Funk et al. 2014).

Temperature data was sourced from the Global Historical Climatology Network version 2 data set and the Climate Anomaly Monitoring System (GHCN CAMS). GHCN CAMS is a high-resolution analysis of global land surface temperatures from 1948 to present, and captures most common temporal-spatial features in the observed climatology and anomaly fields over both regional and global domains (Fan and van den Dool 2008). In addition, we also included altitude (in metres above sea level or 'masl') as an indicator of elevation of the household. This is directly related to the agro-ecological zones, which are mainly classified based on altitude, rainfall, and temperature. Note that, in 2004, altitude readings were not taken, but this was done for the other panels. Altitude is, however, a time-invariant variable and the mean readings for 2007 and 2010 were assumed for 2004.

The variables described in this section can be broadly classified into two: the production function variables, and the exogenous determinants of inefficiency. The descriptive statistics of all the variables used in our analysis are shown in the subsequent subsections. The analysis was done using "R" statistical software.

4.2. Production Function Variables

Farm outputs are captured in kilocalories per acre quantities, with an overall mean of 2739.87 kilocalories. This is a summation of all edible kilocalories produced in the respective farms. Farm inputs considered in the paper include seed, fertilizer, pesticides, herbicides, fungicides, insecticides, and labour. Due to aggregation challenges, seed for different crops, and different types of fertilizers, were converted into Kenyan shillings (Kshs) equivalents. An equivalent conversion was done for fungicides, herbicides, pesticides, and insecticides. These were then summed up as damage control (D/C) inputs. We also converted farm labour into man equivalent units (MEU). The descriptive statistics of the variables used in the first stage of analysis are shown in Table 1 below. In this stage, kilocalories produced per acre was used as the dependent variable.

Table 1. Production Function Output and Input Variables

Year	2004		2007		2010		Full Data	
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
KiloCal/Acre	2,653.26	1,662.21	2,895.99	1,557.78	2,670.37	1,591.38	2,739.87	1,607.78
Seed cost (Ksh.) /Acre	178.32	178.60	221.81	194.85	280.18	195.82	226.77	194.40
Fertilizer cost (Ksh.)/Acre	324.41	334.49	74.44	152.24	390.89	327.59	263.25	315.14
D/C inputs costs (Ksh.) /Acre	90.08	116.63	67.48	86.63	75.14	92.50	77.57	95.60
Labour- Man Equiv. Units/Acre	59.63	28.14	32.52	24.34	31.34	24.00	41.16	28.70
N		1311		1311		1311		3933

4.3. Inefficiency Model Variables

These are further classified into household characteristics and agro-ecological and climatic variables. Details of household characteristics are provided in Table 2. For the three panels combined, mean household age was about 59.28 years. About 43% of the households had persons who were employed and earning a salary outside of farm activities. The education level also varied over the years, with mean education being 6.5 years. In addition, a majority (74%) of the household heads were members of social groups dealing with various social issues. Finally, female-headed households constituted 25% of the total sample.

Table 2. Household Descriptive Statistics by Year

Year	2004		2007		2010		Full Data	
Variable	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Age	59.28	13.34	58.71	13.36	60.43	13.23	59.28	13.34
Age Squared	3,692.15	1,604.57	3,624.96	1,597.73	3,826.51	1,610.97	3,692.15	1,604.57
Female HH Head	0.25	0.43	0.24	0.43	0.27	0.44	0.25	0.43
Salaried household	0.47	0.5	0.45	0.5	0.5	0.5	0.47	0.5
Education (years)	6.46	4.77	6.44	4.74	6.51	4.85	6.46	4.77
Dist. to Main Road	1.06	1.33	0.53	0.84	0.46	0.91	0.68	1.08
Dist. to Extension	5.29	5.79	4.58	5.06	5.4	5.1	5.09	5.34
Group membership	0.76	0.43	0.75	0.43	0.7	0.46	0.74	0.44
N	1311		1311		1311		3933	

Also considered in the analysis was altitude, which was selected to represent agro-ecological factors, while the other variables in Table 3 are climatic factors. Note that altitude is same for all the years as a fixed effect, unlike temperature and precipitation which show variability over time.

Table 3. Agro-Climatic Factors Descriptive Statistics by Year

Year	2004		2007		2010		Full Data	
Statistic	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Altitude (masl)	1,620.80	485.37	1,620.80	485.37	1,620.80	485.37	1,620.80	485.37
Peak Mean Rain	93.32	27.78	112.65	30.75	109.8	37.52	105.26	33.37
Peak Mean Rain sq.	9,479.23	5,352.12	13,634.78	6,760.85	13,463.42	8,424.77	12,192.47	7,218.43
Peak Mean SPEI	-0.78	0.37	0.97	0.4	0.2	0.5	0.16	0.8
Peak Mean SPEI sq.	0.64	0.44	1.12	0.69	0.32	0.29	0.69	0.6
Peak Mean Temp.	20.24	2.53	20.1	2.61	20.61	2.64	20.32	2.6
Peak Mean Temp. sq.	416.12	106.48	410.97	108.82	431.63	112.49	419.57	109.62

5. Results and Discussion

5.1. Farm Productive Efficiency

Efficiency modelling was done using “R” statistical software. The first step in our stochastic frontier analysis was to determine which model – between Cobb-Douglas and translog – best fits our data. This was done by estimating the error components time-invariant stochastic frontier (Batesse and Coelli 1992) using both Cobb-Douglas and translog production functions. The models were unrestricted, i.e., we did not impose any monotonicity or quasi-concavity restrictions. The mean efficiency score for the Translog function was 70.82 while that of the Cobb-Douglas function was 69.61. To determine which model fits our data best, we used the log likelihood ratio test, which is expressed as

$$-2(\ln[L(H_0)] - \ln[L(H_1)] \qquad (17)$$

where $L(H_0)$ and $L(H_1)$ are the values of the log likelihood functions under the null hypothesis (H_0: $\beta_{11} = \beta_{22} = \beta_{12} = 0$) and alternate hypothesis (H_1: $H_0 = False$) (Coelli et al. 1998), i.e., we test that the extra β_s included in the translog model are not different from zero. From our analysis, the log likelihood value for Cobb-Douglas is -4310.3, while that of the translog is -4269.3. The Chi-square value at 10 d.f. is 82.16, against a tabulated value of 18.31 at the 5% level of significance. We therefore rejected the null hypothesis and adapted the translog model as our model of choice. Having established that the translog model was the best fit to our data, we also estimated an error components time-variant translog model (Coelli et al. 1998) whose mean efficiency was 71.66. Figures 1a, 1b and 1c show the distribution of efficiency scores in the three models. The efficiency scores distribution in the three models is similar, with most households in the 80% category.

The estimates from the three models (Cobb-Douglas, translog time-invariant model and translog time-varying model) are shown in Table 4. Note that none of these models has monotonicity or quasi-concavity restrictions imposed. In the three models, the gamma parameter (γ) is significantly different from zero, indicating inefficiency. In addition, in the time-varying efficiency model, the effect of time is negative and significantly different from zero, indicating declining efficiency over time. This is supported by mean efficiencies from the time-varying inefficiency model, which were 75.69, 71.79, and 67.50 for the periods 2004, 2007, and 2010, respectively. Therefore, estimating time-invariant inefficiency models would be erroneous as such a model would assume constancy of efficiency over time.

Figure 1a. Translog Time-Invariant Efficiency

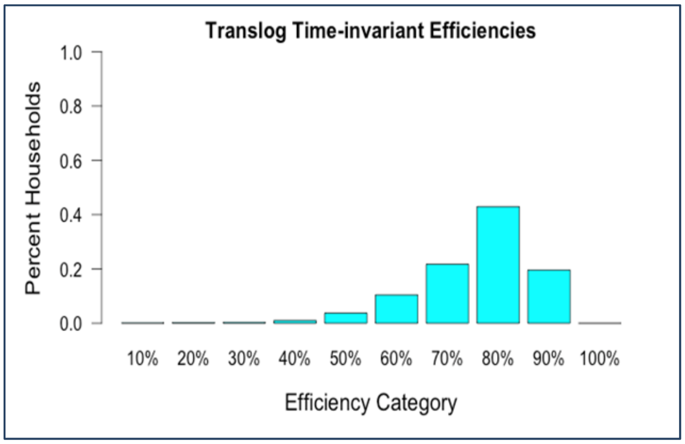


Figure 1b. Cobb-Douglas Time-Invariant Efficiency

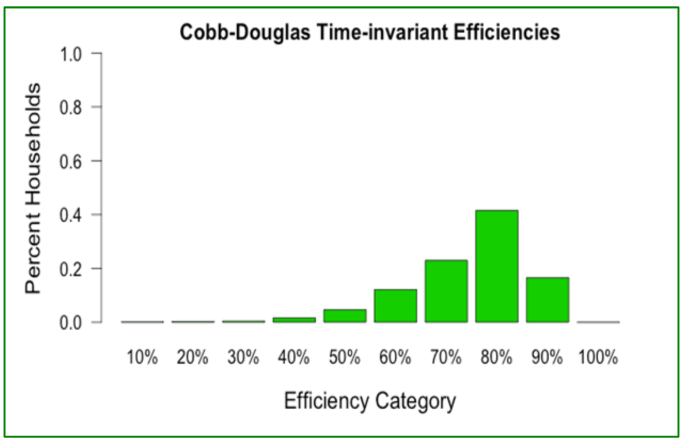


Figure 1c. Translog Time-Varying Efficiency

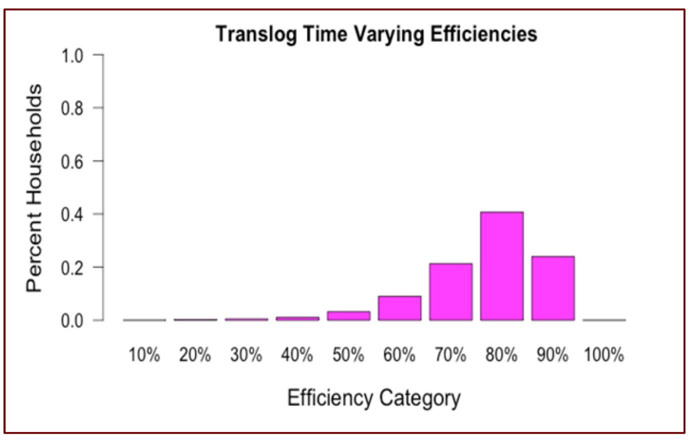


Table 4. Model Estimates of Time-Invariant and Time-Varying Inefficiency Models

Dependent variable: Log Kilo calories			
	Model 1	Model 2	Model 3
VARIABLES	Time-invariant Cobb-Douglas	Time-invariant Translog	Time-varying Translog
Constant	7.6521***	7.9922***	7.9803***
Log Seed	0.0519***	-0.0650**	-0.0749**
Log Fertilizer	0.0154**	0.0538**	0.0732**
Log Other Inputs	0.0334***	0.0626*	0.0532
Log MEU	0.0190	-0.0276	-0.0191
Log Seed Sq.		0.0133**	0.0163***
Log Seed * Log Fert.		0.0036	0.0045
Log Seed * Log Other Inputs		0.0011	0.0019
Log Seed * Log MEU		0.0171**	0.0178**
Log Fert. Sq.		-0.0336***	-0.0349***
Log Fert. * Log Other Inputs		-0.0010	-0.0015
Log Fert. * Log MEU		0.0182***	0.0142**
Log Other Inputs sq.		-0.0061	-0.0058
Log Other Inputs * Log MEU		-0.0041	-0.0024
Log MEU sq.		-0.0328**	-0.0415**
Sigma sq.	0.7102***	0.6788***	0.7618***
Gamma	0.3754***	0.3520***	0.4245***
Time			-0.1913***
Log Likelihood value	-4310.33	-4269.25	-4262.31

5.2. Determinants of Inefficiency

Given evidence of time-varying efficiency, we estimated a technical inefficiency effects model (Coelli et al. 1998) which includes the exogenous determinants of technical inefficiency and also incorporates the time-varying aspect of inefficiency. This was estimated using a translog model with imposed monotonicity restrictions and exogenous determinants of inefficiency, as discussed earlier. Note that quasi-concavity[2] restrictions were not imposed, although, after imposing monotonicity, we tested for quasi-concavity, which held in our model. Table 5 presents the results of this model (with one model using rainfall and temperature as climate variables, and the other using SPEI as the climate variable). Note that a number of exogenous variables influence inefficiency, and variables with negative signs tend to decrease inefficiency while those with positive signs tend to increase inefficiency. For instance, in Model 1, salaried households, education, and membership in a social group are expected to decrease inefficiency. Increased rainfall has a negative sign, and hence it tends to decrease inefficiency. Therefore, households in areas receiving more rainfall are likely to be more efficient. In addition, temperature increases inefficiency and therefore is not favourable for agricultural production. SPEI has an inverted U-shaped relationship with inefficiency, with SPEI increasing inefficiency, but only to a certain point where it assumes a positive relationship, as shown by the negative sign of the quadratic form of SPEI.

[2] Henningsen and Henning (2009) argue that there is not necessarily any technical rationale for production functions to be quasi-concave. They therefore suggest abstaining from imposing quasiconcavity when estimating (frontier) production functions. However, they propose checking for quasiconcavity after the econometric estimation because some standard results of microeconomic theory (e.g., convex input sets) do not hold in the case of non-quasiconcavity.

Table 5. Unrestricted Translog Technical Inefficiency Effects Models

VARIABLES	Model 4	Model 5
	Rainfall and Temperature	SPEI
Constant	8.1870***	8.2170***
Log Seed	-0.0396	-0.0494*
Log Fertilizer	0.0619***	0.0601**
Log Other Inputs	0.0076	0.00001
Log Labour (MEU)	-0.0872*	-0.0835*
Log Seed Squared	0.0178***	0.0185***
Log Seed * Log Fert.	0.0024	0.0024
Log Seed * Log Other Inputs	0.0013	0.0021
Log Seed * Log MEU	0.0051	0.0071
Log Fert. Sq.	-0.0277***	-0.0249***
Log Fert. * Log Other Inputs	-0.0026	-0.0022
Log Fert. * Log MEU	0.0131**	0.0124**
Log Other Inputs Sq.	0.0062	0.0034
Log Other Inputs * Log MEU	0.0017	0.0034
Log MEU Sq.	0.0042	-0.0023
Age	2.1597***	-22.8880**
Age Sq.	-0.0112***	0.1960**
House Hold Size	1.2767**	7.8233**
Female Headed HH	15.4040***	64.7850**
Salaried H. Hold	-1.2112***	18.6000**
Educ. In years	-0.8264***	-7.9670**
Dist. To Road	3.7975*	10.6700**
Dist. To Extn.	1.6526**	9.9675**
Member Soc. Grp	-15.3110**	-111.2400**
Elevation (masl)	0.0090***	-0.0123*
Mean Rain	-95.2090**	-
Mean Temp	40.2020***	-
SPEI	-	22.0610**
SPEI sq.	-	-328.7300**
Sigma Sq.	116.4900**	462.4900**
Gamma	0.9987***	0.9997***

The same interpretation can be given to other non-climatic variables. The mean efficiency scores for these two models were 64.22 and 64.24 respectively and the distribution of these efficiencies are as shown in Figures 2a and 2b. The distributions of these models are similar and have more spread compared to the time-invariant and time-varying models.

Figure 2a. Unrestricted Inefficiency Effect Model 1

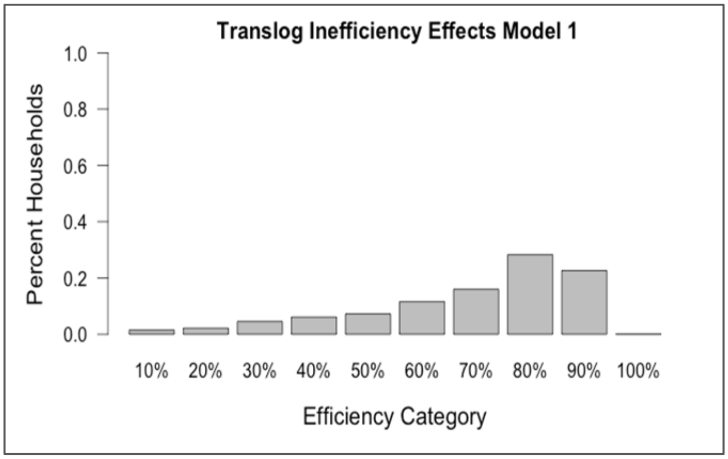
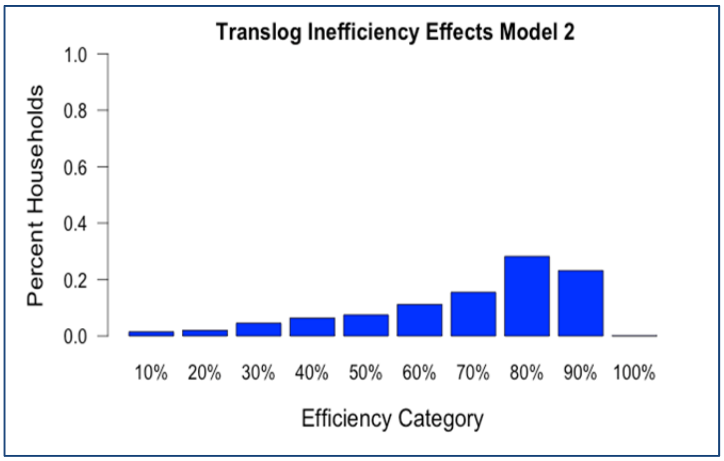


Figure 2b. Unrestricted Inefficiency Effect Model 2



5.3. Imposing Monotonicity in Technical Inefficiency Effects Model

Monotonicity was imposed on these two models and the estimates from the resultant monotonic models are shown in Table 6. The coefficient "Fitted model" in both models comprises the combined effect of the coefficients of the production model. The other variables remain the same as in the unrestricted models. The climate variability (weather) factors have the correct signs in the monotonic models, although they are not significant. Most of the other variables in Model 2 (SPEI) are significant, and we can draw policy conclusions from this model.

Table 6. Technical Inefficiency Effects Models with Imposed Monotonicity

VARIABLES	Model 6 (Rain and Temperature)	Model 7 (SPEI)
Constant	0.12764	0.1379
Fitted Model	0.98556***	0.9840***
Age	1.87145	-4.4240***
Age sq.	-0.00965	0.0382***
Household Size	0.90734*	1.6783**
Female Headed HH	14.30459	11.2730***
Salaried H. Hold	-2.17408	2.4380
Educ. In years	-0.50815	-1.4309**
Dist. To Road	3.07179	2.5985*
Dist. To Extn.	1.47502	1.9713**
Member Soc. Grp	-10.99495	-21.6590**
Elevation (masl)	0.00332***	-0.0045**
Mean Rain	-81.86291	-
Mean Temp	37.11204	-
SPEI	-	0.6006
SPEI sq.	-	-71.4850***
Sigma Sq.	99.43130	93.4540***
Gamma	0.99849***	0.9984***

The adjusted coefficients of the models with imposed monotonicity are as shown in Table 7. These are compared with those from the unrestricted models (without monotonicity). These coefficients are different and therefore making inferences or policies from non-monotonic production functions may result in different interpretations.

Table 7. Adjusted Coefficients for Models with Imposed Monotonicity and Unrestricted Models

VARIABLES	Models with Monotonicity		Unrestricted Models	
	Model 8	Model 9	Model 10	Model 11
	Rainfall and Temperature	SPEI	Rainfall and Temperature	SPEI
Constant	7.9123	7.8971	8.1870	8.2170
Log Seed	0.0220	0.0229	-0.0396	-0.0494
Log Fertilizer	0.0445	0.0507	0.0619	0.0601
Log Other Inputs	0.0064	0.0048	0.0076	0.00001
Log MEU	0.0022	0.0006	-0.0872	-0.0835
Log Seed Sq.	0.0063	0.0064	0.0178	0.0185
Log Seed * Log Fert.	0.0000	0.0001	0.0024	0.0024
Log Seed * Log Other Inputs	0.0014	0.0019	0.0013	0.0021
Log Seed * Log MEU	0.0004	0.0001	0.0051	0.0071
Log Fert. Sq.	-0.0062	-0.0068	-0.0277	-0.0249
Log Fert. * Log Other Inputs	-0.0002	-0.0006	-0.0026	-0.0022
Log Fert. * Log MEU	0.0002	0.0002	0.0131	0.0124
Log Other Inputs sq.	-0.0004	0.0004	0.0062	0.0034
Log Other Inputs * Log MEU	0.0034	0.0027	0.0017	0.0034
Log MEU sq.	-0.0003	-0.0001	0.0042	-0.0023

The efficiency scores of these models are as shown in Figures 3a and 3b. The mean efficiencies are 63.67 and 63.91, respectively. These efficiencies are marginally higher than those of the unrestricted models, but statistically not different. The dispersion of the scores is also very similar to that of the unrestricted models. The implication is that adjusting for monotonicity gives the model grounding in economic theory and the results obtained from the model are more reliable than those from unrestricted models. The efficiency scores and dispersion are not very different from the unrestricted model, but the coefficients are different.

Figure 3a. Restricted Inefficiency Effects Model 1

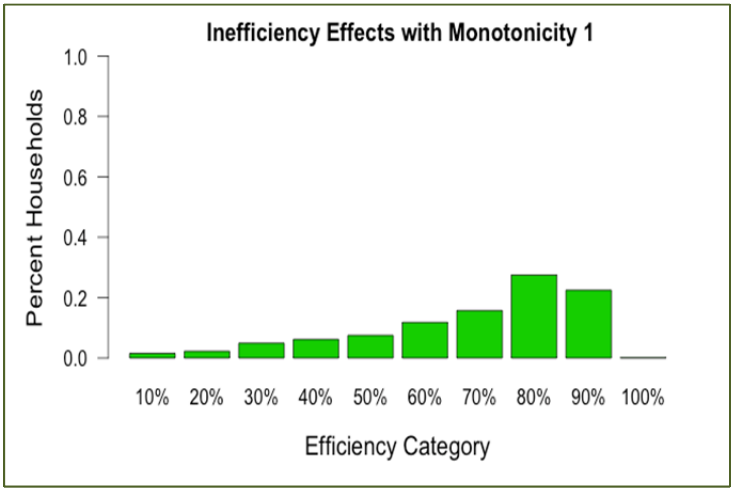
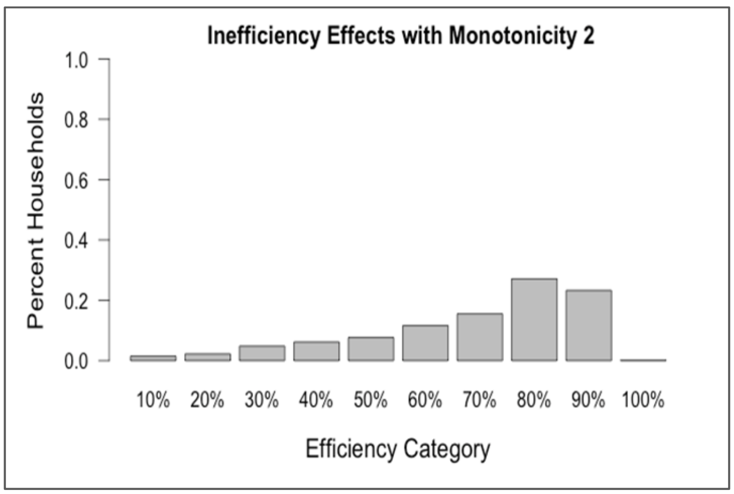


Figure 3b. Restricted Inefficiency Effects Model 2



6. Conclusions and Recommendations

In this analysis, we have demonstrated how efficiency can be estimated from a monotonic translog frontier model, an approach that has not been applied in most efficiency studies. The data used were panel data from three waves, and exogenous determinants of efficiency were included in the model. From the results, farmers in Kenya are, on average, 63% efficient and thus could expand output by about 37% and still use the same level of inputs. Therefore, even without utilizing more seed, fertilizer, labour and damage control inputs, it is feasible for farmers to significantly expand farm level production, and hence efficiency. However, this will be determined by a number of exogenous variables, which are climatic, agro-ecological, and household factors. Climatic variability factors such as higher rainfall tend to decrease inefficiency, while increased temperature increases inefficiency. From a policy perspective, these can be addressed though planned or autonomous adaptation of technologies such as improved seed and irrigation.

Some household factors do influence inefficiency, either positively or negatively. While some of them can be addressed through policy, others are more or less fixed in the short term. The quasi-fixed factors such as education tend to decrease inefficiency, while household size increases inefficiency. Addressing these from a policy perspective would require long-term policy decisions such as investments in education and family planning. Age of the household head tends to decrease inefficiency, and it is expected that older, experienced farmers are more efficient. Targeting therefore needs to be focused on the younger farmers, so as to train them on climate change and variability adaptation techniques, and also good agronomic practices. In addition, some traditional knowledge can be borrowed from the older farmers in tackling climate change and variability. Membership in farmer groups is a measure of social capital and decreases inefficiency. Farmer groups and other groups could therefore be a good entry point of introducing adaptation and other farming techniques. Distance to extension agents tends to increase inefficiency, and farmers who are farther from extension agents tend to be more inefficient. This could be explained by their limited access to extension due to distance. This can be addressed by increasing extension funding from national and county governments so as to increase the number of extension agents at county levels. County governments can also facilitate extension agents and ensure that public extension services are reaching all households in the respective counties.

References

Aigner, D.J., C.A.K. Lovell and P. Schmidt. 1977. Formulation and Estimation of Stochastic Frontier Production Functions. *Journal of Econometrics* 6(1): 21-37.

Alemu, B.A., E.-A. Nuppenau and H. Boland. 2009. Technical Efficiency of Farming Systems across Agro-ecological Zones in Ethiopia: An Application of Stochastic Frontier Analysis. Justus-Liebig University of Giessen, Germany, Department of Environmental Management and Market Research.

Ali, M., and J.C. Flinn. 1989. Profit Efficiency among Basmati Rice Producers in Pakistan Punjab. *American Journal of Agricultural Economics* 71(2): 303-310.

Asada, H., and J. Matsumoto. 2009. Effects of Rainfall Variation on Rice Production in the Ganges–Brahmaputra Basin. *Climate Resource* 38: 249-260.

Battese, G.E., and G.S. Corra. 1977. Estimation of a Production Frontier Model: With Application to the Pastoral Zone off Eastern Australia. *Australian Journal of Agricultural Economics* 21(3): 169-179.

Battese, G.E., and T.J. Coelli. 1988. Prediction of Firm-level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data. *Journal of Econometrics* 38: 387-399.

Battese, G.E., and T.J. Coelli. 1992. Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis* 3: 153-169.

Battese, G.E., and T.J. Coelli. 1995. A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data. *Empirical Economics* 20(1): 325-332.

Beguería, S., and S.M. Vicente-Serrano. 2013. SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index, (R package version 1.6 edn.).

Bravo-Ureta, B.E., and L. Rieger. 1991. Dairy Farm Efficiency Measurement Using Stochastic Frontiers and Neoclassical Duality. *American Journal of Agricultural Economics* 73(2): 421-428.

Brazdik, F. 2006. Non-parametric Analysis of Technical Efficiency: Factors Affecting Efficiency of West Java Rice Farms. CERGE-EI Working Paper 286. The Center for Economic Research and Graduate Education-Economic Institute, Prague.

Cherchye, L., T. Kuosmanen and T. Post. 2000. What is the Economic Meaning of FDH? A Reply to Thrall. *Journal of Productivity Analysis* 13(3): 263-267.

Coelli, T., D.S.P. Rao and G.E. Battese. 2005. *An Introduction to Efficiency and Productivity Analysis*. Boston, MA: Kluwer Academic Publishers.

Coelli, T.J., D.S. Prasada Rao and G. Battese. 1998. *An Introduction to Efficiency and Productivity Analysis.* Norwell, MA: Kluwer Academic Publishers.

Deressa, T.T., R.M. Hassan, C. Ringler, T. Alemu and M. Yesuf, 2009. Determinants of Farmers' Choice of Adaptation Methods to Climate Change in the Nile Basin of Ethiopia. *Global Environmental Change* 19(2)**:** 248-255.

Diewert, E.W. 1974. Applications of Duality Theory. In *Frontiers of Quantitative Economics Vol. II*, edited by M. Intriligator and D. Kendrick. Amsterdam: North-Holland Publishing, 106-71.

Emrouznejad, A., and V. Podinovski. 2004. *Data Envelopment Analysis and Performance Management*. UK: Warwick Print.

Fan, Y., and H. Van den Dool. 2008. A Global Monthly Land Surface Air Temperature Analysis for 1948–Present. *Journal of Geophysical Research: Atmospheres (1984–2012)* 113(D1).

Farrell, M.J. 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society Series A, General* 120(3): 253-281.

Funk, C., A. Hoell, S. Shukla, I. Bladé, B. Liebmann, J.B. Roberts, F.R. Robertson and G. Husak. 2014. Predicting East African Spring Droughts Using Pacific and Indian Ocean Sea Surface Temperature Indices. *Hydrology and Earth System Sciences* 18: 4965-4978.

Gbetibouo, G.A., and R.M. Hassan. 2005. Measuring the Economic Impact of Climate Change on Major South African Field Crops: A Ricardian Approach. *Global and Planetary Change* 47(2-4): 143-152.

Greene, W.H. 1980. Maximum Likelihood Estimation of Econometric Frontier Functions. *Journal of Econometrics* 13(1): 27-56.

Greene, W.H. 1990. A Gamma-distributed Stochastic Frontier Model. *Journal of Econometrics* 46(1-2): 141-164.

Henderson, B., and R. Kingwell. 2005. Rainfall and Farm Efficiency Measurement for Broadacre Agriculture in South-Western Australia. *Australasian Agribusiness Review* 13(14). ISSN 1442-6951.

Henningsen, A., and C. Henning. 2009. Imposing Regional Monotonicity on Translog Stochastic Production Frontiers with a Simple Three-Step Procedure. *Journal of Productivity Analysis* 32(3): 217-229.

Huang, C.J., and J.-T. Liu. 1994. Estimation of a Non-neutral Stochastic Frontier Production Function. *Journal of Productivity Analysis* 5(1): 171-180.

IPCC. 2007. *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the IPPC*. Cambridge, UK: Cambridge University Press, 869-883.

Jondrow, J., C.A.K. Lovell, I.S. Materov and P. Schmidt (JLMS). 1982. On the Estimation of Technical Efficiency in the Stochastic Frontier Production Function Model. *Journal of Econometrics* 19(2/3): 233-238.

Kabara, M., and J. Kabubo-Mariara. 2011. Global Warming in the 21st Century: The Impact on Agricultural Production in Kenya. In *Global Warming in the 21st Century*, edited by J.M. Cossia. New York: Nova Science Publishers, 199-214.

Kabubo-Mariara, J., and M. Kabara. 2015. Climate Change and Food Security in Kenya. Environment for Development Discussion Paper Series, EfD DP 15-05.

Kabubo-Mariara, J., R. Mulwa and S. Di Falco. 2016. The Impact of Climate Change on Food Calorie Production and Nutritional Poverty: Evidence from Kenya. Environment for Development Discussion Paper Series, EfD DP 16-26.

Kabubo-Mariara, J., R. Mulwa and S. Di Falco. 2017. Adaptation to Climate Change and Its Implications for Food Security among Smallholder Farmers in Kenya. Environment for Development Discussion Paper Series 17-05.

Kalirajan, K.P. 1990. On Measuring Economic Efficiency. *Journal of Applied Econometrics* 5(1): 75-85.

Kibaara, B.W. 2005. *Technical Efficiency in Kenyan's Maize Production: An Application of the Stochastic Frontier Approach*. USA: Colorado State University.

Kirimi, L., and S.M. Swinton. 2004. Estimating Cost Efficiency among Maize Producers in Kenya and Uganda. Selected Paper Prepared for Presentation at the American Agricultural Economics Association Annual Meeting, Denver, Colorado, 1-4.

Koebel, B., M. Falk and F. Laisney. 2003. Imposing and Testing Curvature Conditions on a Box-Cox Cost Function. *Journal of Business Economic Statistics* 21(2): 319-335.

Kumar, K.K., K.R. Kumar, R.G. Ashrit, N.R. Deshpande and J.W. Hansen. 2004. Climate Impacts on Indian Agriculture. *International Journal of Climatology* 24: 1375-1393.

Kumbhakar, S.C., and C.A.K. Lovell. 2000. *Stochastic Frontier Analysis*. Cambridge, UK: Cambridge University Press.

Kumbhakar, S.C., S. Ghosh and J.T. McGuckin. 1991. A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in US Dairy Farms. *Journal of Business and Economic Statistics* 9(1): 279-286.

Lobell, D.B., and M.B. Burke. 2008. Why Are Agricultural Impacts of Climate Change So Uncertain? The Importance of Temperature Relative to Precipitation. *Environmental Research Letters* 3: 1-8.

Makki, M.F., Y. Ferrianta and R. Suslinawati. 2012. Impacts of Climate Change on Productivity and Efficiency Paddy Farms: Empirical Evidence on Tidal Swamp Land South Kalimantan Province – Indonesia. *Journal of Economics and Sustainable Development* 3(14): 66-72.

Meeusen, W., and J. van den Broeck. 1977. Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review* 18(2): 435-444.

Mendelsohn, R., W. Nordhaus and D. Shaw. 1994. The Impact of Global Warming on Agriculture: A Ricardian Analysis. *American Economic Review* 84: 753-771.

Mkhabela, T., J. Piesse, C. Thirtle and N. Vink. 2010. Modelling Efficiency with Farm Produced Inputs: Dairying in KwaZulu-Natal, South Africa. *Agrekon* 49(1): 102-121.

Mulwa, R., and J. Kabubo-Mariara 2017. Influence of Agro-climatic Factors and Adaptation Strategies on Productive Efficiency in Kenya: A DEA Semi-parametric Framework. Environment for Development Discussion Paper Series *(Forthcoming)*

Mulwa, R., A. Emrouznejad, and L. Muhammad. 2009. Economic Efficiency of Smallholder Maize Producers in Western Kenya: A DEA Meta-Frontier Analysis, *International Journal of Operational Research* 4(3): 250-267.

Mulwa, R., K.P.C. Rao, S. Gummadi, and M. Kilavi. 2016. Effects of Climate Change on Agricultural Households' Welfare in Kenya. *Climate Research* (in Press).

Mundlak, Y. 1961. Empirical Production Function Free of Management Bias. *Journal of Farm Economics* 43(1): 44-56.

Nyagaka, D.O., G.A. Obare, J.M. Omiti, and W. Nguyo. 2010. Technical Efficiency in Resource Use: Evidence from Smallholder Irish Potato Farmers in Nyandarua North District, Kenya. *African Journal of Agricultural Research* 5(1): 1179-1186.

O'Donnell, C.J., and T.J. Coelli. 2005. A Bayesian Approach to Imposing Curvature on Distance Functions. *Journal of Econometrics* 126(2): 493-523.

Ogada, M.J., D. Muchai, G. Mwabu and M. Mathenge. 2014. Technical Efficiency of Kenya's Smallholder Food Crop Farmers: Do Environmental Factors Matter? *Environment Development and Sustainability* 16(5): 1065-1076.

Oyekale S. A., (2012). Impact of Climate Change on Cocoa Agriculture and Technical Efficiency of Cocoa Farmers in South-West Nigeria. *Journal Human Ecology,* Vol.40 (2): 143-148

Pitt, M.M., and L.-F. Lee. 1981. The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics* 9(1): 43-64.

Reifschneider, D., and R. Stevenson. 1991. Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency. *International Economic Review* 32(1): 715-723.

Sauer, J., K. Frohberg and H. Hockmann. 2006. Stochastic Efficiency Measurement: The Curse of Theoretical Consistency. *Journal of Applied Economics* 9(1): 139-165.

Schmidt, P., and R.C. Sickles. 1984. Production Frontiers and Panel Data. *Journal of Business and Economic Statistics* 2(2): 367-374.

Sherlund, S.M., C.B. Barrett and A.A. Adesina. 2002. Smallholder Technical Efficiency Controlling for Environmental Production Conditions. *Journal of Development Economics* 69: 85-101.

Stevenson, R. 1980. Likelihood Functions for Generalized Stochastic Frontier Estimation. *Journal of Econometrics* 13(1): 58-66.

Thrall, R.M. 1999. What is the Economic Meaning of FDH? *Journal of Productivity Analysis* 11(3): 243-250.

Wang, H.-J., and P. Schmidt. 2002. One-Step and Two-step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels. *Journal of Productivity Analysis* 18: 129-144.

Wollni, M., and B. Brummer. 2012. Productive Efficiency of Speciality and Conventional Coffee Farmers in Costa Rica: Accounting for Technological Heterogeneity and Self-selection. *Food Policy* 37(1): 67-76.

Wooldridge, J.M. 2010. *Econometric Analysis of Cross Section and Panel Data,* 2nd Edition. Cambridge, Massachusetts: MIT Press.

Appendix

Appendix 1: Variables Used in the Model

Production function		Exogenous inefficiency determinants		
Dependent variable				
	Production variables	Household Variables	Climate Variables	Agro-Ecological variable
Kilocalories/Acre	Seed cost	Age	Peak Mean rainfall	Altitude (masl)
	Fertilizer cost	Age squared	Peak Mean rainfall sq.	
	Damage control inputs costs	Household size	Peak Mean SPEI	
	Labour (MEU)	Female-headed	Peak Mean SPEI sq.	
	Seed cost	Salaried households	Peak Mean temp.	
		Education	Peak Mean temp. sq.	
		Distance to main road		
		Distance to extension		
		Member of social group		