Hoang, Viet-Ngu & Nguyen, Trung Thanh (2013) Analysis of environmental efficiency variation : a materials balance approach. *Ecological Economics*, *86*(1), pp. 37-46.

ANALYSIS OF ENVIRONMENTAL EFFICIENCY VARIATION: A MATERIALS BALANCE APPROACH

Abstract

Recent literature has argued that environmental efficiency (EE), which is built on the materials balance (MB) principle, is more suitable than other EE measures in situations where the law of mass conversation regulates production processes. In addition, the MB-based EE method is particularly useful in analysing possible trade-offs between cost and environmental performance. Identifying determinants of MB-based EE can provide useful information to decision makers but there are very few empirical investigations into this issue. This article proposes the use of data envelopment analysis and stochastic frontier analysis techniques to analyse variation in MB-based EE. Specifically, the article develops a stochastic nutrient frontier and nutrient inefficiency model to analyse determinants of MB-based EE. The empirical study applies both techniques to investigate MB-based EE of 96 rice farms in South Korea. The size of land, fertiliser consumption intensity, cost allocative efficiency, and the share of owned land out of total land are found to be correlated with MB-based EE. The results confirm the presence of a trade-off between MB-based EE and cost allocative efficiency and this finding, favouring policy interventions to help farms simultaneously achieve cost efficiency and MP-based EE.

Keywords: environmental efficiency, materials balance, nutrient efficiency, nutrient stochastic frontier, single-bootstrap truncated regression

1. Introduction

There are two important components in any type of empirical environmental efficiency (EE) analysis, particularly in agricultural production (Reinhard et al., 2002). The first component estimates EE scores and variation in the EE scores across farms, the second identifies determinants of such variation. For the first component, several approaches to measuring EE exist (for an overview see for example Tyteca 1996 and Callens and Tyteca1999). Recent literature favours the use of those EE measures which are based on the balances of materials, particularly in an agricultural sector (hereafter called MB-based EE) (Coelli et al., 2007; Hoang and Coelli, 2011; Lauwers, 2009). MB-based EE measures are preferred because the materials balance principle (MBP) regulates the transformation of materials in such closed systems of agricultural production; hence EE measures, in order to be reliable, should be adjusted to be consistent with the MBP. Moreover, the MB-based approach can lead to a more diversified analysis of EE and facilitate analysis of trade-offs between the economic and environmental performance of a given production technology (Lauwers, 2009; Van Meensel et al., 2010).

The MB-based approach has been applied in analysing the EE of several types of decision-making units (DMUs) in crop and livestock production in which the balances of nutrients such as nitrogen (N) and phosphorous (P) are considered as polluting emissions. Reinhard and Thijssen (2002) analysed Dutch dairy farms using a stochastic frontier analysis (SFA) technique. Coelli et al. (2007) investigated the environmental performance of 117 pig finishing farms in Belgium using a data envelopment analysis (DEA) technique. Van Meensel et al. (2010) applied both DEA and SFA techniques to the same data set used in Coelli et al. (2007) to analyse trade-offs between EE and economic efficiency. Hoang and

Coelli (2011) and Hoang and Alauddin (2012) studied crop and livestock production in developed countries using the DEA technique. Nguyen et al. (2012) investigated the environmental performance of rice farms in South Korea. These studies found high variation in MB-based EE across decision-making units (DMUs) (i.e., countries and farms). For example, Nguyen et al. (2012) reported remarkably high variation of MB-based EE across 196 rice farms (e.g., a mean EE score: 0.309, the range: 0.055 to 1, and standard deviation: 0.179) (Nguyen et al., 2012).

With respect to the second component, the identification of determinants of variation, this type of analysis can provide decision makers with useful information about how to improve EE. Several analytical frameworks (for example two-stage DEA models or single-stage SFA models) to analyse efficiency drivers have been well developed and widely used in empirical studies (Battese and Coelli, 1995; Coelli et al., 2005; Greene, 2005; Simar and Wilson, 2007). Researchers have used these frameworks to investigate drivers of EE variation. Reinhard et al. (2002) appear to be one of the most cited empirical studies that investigate the determinants of EE variation in the context of agricultural production; however, this study uses an EE model that is not adjusted for the MBP.

However, none of previous empirical studies of the MB-based EE approach performed the second component of the analysis. Hence, it is desirable to assess critically whether the existing analytical frameworks of analysing EE determinants can be appropriate in the context of MB-based EE analysis. The present article aims to fill this gap by using bootstrap truncated two-stage DEA models proposed by Simar and Wilson (2007) and estimating the stochastic nutrient frontier following the stochastic frontier model of Battese and Coelli (1995). Empirical applications of these models into a data set of rice farms in South Korea also illustrated the possibility of conducting a statistical hypothesis test for trade-offs between economic and environmental performance.

The remainer of the article is structured as follows. Section 2 provides a brief literature review on various approaches to measuring EE. Section 3 provides a mathematical illustration of the shortcoming of the EAPE model in relation to the MBP. Section 4 reviews the MB-based EE method and discusses potential uses of this method for trade-off and policy analysis. Section 5 introduces the use of SFA and DEA techniques to analyse variation in the MB-based EE. Section 6 presents an empirical analysis of rice farms in South Korea. Section 7 concludes the article.

2. Main approaches to measuring environmental efficiency: a literature review

Lauwers (2009) provides a review of three general groups of models used to measure EE: the environmentally adjusted production efficiency, the frontier eco-efficiency and the MB-based models. The environmentally adjusted production efficiency (EAPE) uses the production frontier to analyse a relationship between inputs and outputs. In EAPE's models, pollution is viewed as either environmentally detrimental inputs or undesirable outputs. Adding pollution as an extra input or output in conventional production models, technical efficiency (TE) measures can be estimated with input-oriented, output-orientated frameworks, or with hyperbolic or directional distance functions (Chung et al., 1997; Färe et al., 2007; Färe et al., 1996; Reinhard et al., 2002). An input-orientated framework minimises inputs given fixed output quantities. An output-orientated framework maximises outputs with fixed input quantities. The hyperbolic and directional distance functions allow the simultaneous expansion of outputs and the contraction of inputs. The proponents of these methods argue that these models credit farms for the contraction of pollution; therefore, TE can be interpreted as EE.

The frontier eco-efficiency (FEE) uses the frontier framework to model relationships between economic and ecological outcomes to derive eco-efficiency measures (Callens and Tyteca, 1999; Tyteca, 1999). The eco-efficiency measures relate the economic value of outputs to the environmental pressures involved in production processes (Picazo-Tadeo et al., 2012). Several empirical studies have applied this approach (Kortelainen, 2008; Kuosmanen and Kortelainen, 2005; Picazo-Tadeo et al., 2011). These applications can be seen as the frontier operationalisation of the eco-efficiency concept in the analysis of multidimensional sustainability (Lauwers, 2009). For example, Picazo-Tadeo et al. (2011), using a data set of 117 crop farms in Spain, assessed the opportunities of reducing five environmental pressures (tendency towards monoculture that has potential impacts on biodiversity, N balance, P balance, energy balance, and pesticide risks), given the value added of crop outputs.

There is a methodological distinction between EAPE and FEE models. The EAPE models are based on the conventional production relationship between inputs and outputs while the FEE models are grounded on a hypothesised relationship between economic values of outputs and environmental pressures. Often they are used in different research contexts. The primary use of the EAPE approach is to adjust efficiency measures to account for environmental pollution in the paradigm of costly environmental regulation. In this paradigm, efficiency analysis methods implicitly suppose that efficiency improvements imply cost reduction (Lauwers, 2009). The FEE approach is used mainly to provide relative assessments among DMUs in terms of environmental performance where there are many types of environmental pressures caused by production and consumption activities.

The third approach to measuring EE involves the use of the MB-based models firstly proposed by Coelli et al. (2007). The MB-based models view pollution as the balance of materials and attempt to minimise this balance. The MB-based EE measures are defined as the technically feasible minimum materials balance to the currently observed materials balance. The MB-based models are distinct from the EAPE and FEE methods because the materials balance does not appear as either an input/output in EAPE models or an indicator of environmental pressures in FEE models.

Note that the MB-based and EAPE models are grounded on the same production relationship between inputs and outputs; hence they are very useful in analysing economic-environmental trade-offs faced by DMUs (Lauwers, 2009; Van Meensel et al., 2010). However, the MB-based models are more suitable in situations where the MBP regulates the transformation of materials in production processes (Hoang and Alauddin, 2012; Hoang and Coelli, 2011; Nguyen et al., 2012).[1] The MB-based models are preferred because given the existing construction of EAPE models, measuring environmental inefficiency as the degree to which pollution (i.e., the materials balance) can be reduced with traditional inputs and outputs held

[1] For example, in rice production not all N and P in seed, chemical fertilisers, organic fertilisers and land are transformed into rice outputs. In fact, N and P balances, defined as the differences of the total amounts of N and P in inputs and of the total amounts of N and P in outputs, will go to water and atmospheric environments. Scientifically, these balances have been identified as the main cause of eutrophication in lake, river, and ocean water systems (Smith et al., 1999). Therefore, in rice production, the balance of nutrients can be considered as potential polluting agents.

constant is mathematically infeasible (Coelli et al., 2007; Hoang and Coelli, 2011; Lauwers, 2009). To provide further evidence of this shortcoming of the EAPE models, the next section investigates the model of Reinhard et al. (2002) in which emission is modelled as an input.

3. A major shortcoming of the EAPE models: a simple mathematical illustration

Consider the situation where farms produce a vector of M outputs, $\mathbf{q} \in \Re_{+}^{M}$, using a vector of K inputs, $\mathbf{x} \in \Re_{+}^{K}$. The production activity also produces an emission of polluting substances. The amount of emission is defined by the balance of nutrients:

$$(1) \qquad u = \mathbf{ax} - \mathbf{bq}$$

where **a** and **b** are the vectors representing nutrient contents of inputs and outputs. Some inputs, such as labour and machinery, could have zero contents of nutrients, suggesting that vectors **a** may include zero values.

The MBP applies to individual flows of nutrients (e.g., N or P). In situations where there are many types of nutrients involved, one can use weights that reflect the polluting power of different nutrients in calculating the aggregate nutrient balance. For example, N and P are two main causes of eutrophication (i.e. oxygen depletions caused by excessive nutrient-induced increases in the production of organic matter) in water systems (Howarth et al., 2000). The analysis of eutrophication requires the use of a particular set of weights that reflect the eutrophying power of N and P in the context of a specific water system such that their aggregate effects can be analysed in empirical studies. Given an appropriate choice of N:P weights, the aggregate balance of nutrients can be calculated. Note that the eutrophying powers of N and P depend on the nature of the systems. Systems such as lakes and rivers tend to be limited more by P than N. Furthermore, the over-enrichment of P results in a more damaging effect than the over-enrichment of N. In contrast, N is more commonly the key limiting nutrient of marine waters; thus, N levels have a greater eutrophying power in salt water systems than do the P levels. Hence, it is important that weights for N and P be carefully determined (Hoang and Coelli, 2011).

Given technology with observed outputs (**q**) and conventional inputs (**x**), one can follow the EAPE model (e.g., Reinhard et al., 2002) and define EE as the ratio of minimum feasible to observed use of an environmentally detrimental input. This EAPE's EE measure involves trying to find the largest scalar, λ, such that the scaled vector ($\mathbf{q}, \mathbf{x}, \mathbf{u}/\lambda$) is within the feasible production set. In this context, more environmentally efficient farms use less **u**. Applying this scaling to the MBP results in:

(2) $$\mathbf{u} / \lambda = \mathbf{ax} - \mathbf{bq}$$

Combining (1) and (2), we have:

(3) $$\mathbf{u} = \mathbf{u} / \lambda$$

Thus the only solution to (3) is $\lambda = 1$. This solution, however, refers only to farms that are on the production frontier. Farms that lie below the production frontier are not mathematically feasible to the constraint of the MBP. This restriction is an undesirable feature.

In review the MB-based models is more suitable for empirical studies in situations where researchers believe that the MBP should be applied. The next section recalls the essence of the MB-based model of Coelli et al. (2007) in the context of agricultural production where we are concerned the potential polluting effects caused by the balance of nutrients.

4. The MB-based environmental efficiency measure

The balance of nutrients in (1) is defined as an environmental pressure. When outputs are fixed, the nutrient balance is minimised when the total amount of nutrients in inputs ($\mathrm{NC} = \mathbf{a'x}$) is minimised. Instead of minimising inputs (**x**), we minimise the total amount of nutrients contained in **x**. In the input-orientated framework, this approach involves the following optimisation problem:

$$(4) \qquad \mathrm{NC}(\mathbf{q},\mathbf{a}) = \min_{\mathbf{x}}\{\mathbf{a'x} | \langle \mathbf{x},\mathbf{q} \rangle \in \mathrm{T}\}$$

where T is a feasible production set that is defined as:

$$(5) \qquad \mathrm{T} = \{(\mathbf{q},\mathbf{x}) : \mathbf{x} \text{ can produce } \mathbf{q}\}$$

$\mathrm{NC_{NE}}$ is a solution to (4) and the input vector involved in this minimum nutrient amount is $\mathbf{x}_{\mathrm{NE}}$ when $\mathrm{NC_{NE}} = \mathbf{a'x}_{\mathrm{NE}}$. The input-orientated approach defines nutrient-orientated EE (hereinafter named NE) as the ratio of the minimum nutrient amount to the observed nutrient amount:

$$(6) \qquad \mathrm{NE} = \frac{\mathrm{NC_{NE}}}{\mathrm{NC}} = \frac{\mathbf{a'x}_{\mathrm{NE}}}{\mathbf{a'x}}$$

Input-orientated technical efficiency (TE) is defined as:

$$\text{(7)} \quad \mathrm{TE}(\mathbf{q},\mathbf{x}) = \min_{\theta}\{\theta \mid \langle \theta\mathbf{x},\mathbf{q}\rangle \in \mathrm{T}\}$$

where θ is a scalar taking a value between zero and one. TE addresses the question of the proportional reduction of input quantities while producing a given level of output quantities. Equation (7) has a solution $\mathbf{x}_{\mathrm{TE}}$ that is technically efficient with the total amount of nutrients, $\mathrm{NC}_{\mathrm{TE}} = \mathbf{a}'\mathbf{x}_{\mathrm{TE}}$. TE can also be written as:

$$\text{(8)} \quad \mathrm{TE} = \theta = \frac{\mathbf{a}'\mathbf{x}_{\mathrm{TE}}}{\mathbf{a}'\mathbf{x}} = \frac{\mathrm{NC}_{\mathrm{TE}}}{\mathrm{NC}}$$

NE can be decomposed into TE, and the input-orientated nutrient allocative efficiency is as follows:

$$\text{(9)} \quad \mathrm{NE} = \frac{\mathrm{NC}_{\mathrm{NE}}}{\mathrm{NC}} = \frac{\mathbf{a}'\mathbf{x}_{\mathrm{INE}}}{\mathbf{a}'\mathbf{x}} = \frac{\mathbf{a}'\mathbf{x}_{\mathrm{TE}}}{\mathbf{a}'\mathbf{x}} \times \frac{\mathbf{a}'\mathbf{x}_{\mathrm{NE}}}{\mathbf{a}'\mathbf{x}_{\mathrm{TE}}} = \frac{\mathrm{NC}_{\mathrm{TE}}}{\mathrm{NC}} \times \frac{\mathrm{NC}_{\mathrm{NE}}}{\mathrm{NC}_{\mathrm{TE}}} = \mathrm{TE} \times \mathrm{NAE}$$

where

$$\text{(10)} \quad \mathrm{NAE} = \frac{\mathbf{a}'\mathbf{x}_{\mathrm{NE}}}{\mathbf{a}'\mathbf{x}_{\mathrm{TE}}}$$

TE can be estimated by a standard input-orientated framework, whereas NE can be estimated following a procedure similar to estimating cost efficiency in which the vector of nutrient contents of inputs is used instead of prices. Residually, NAE can be estimated as a ratio of NE to TE. The decomposition in (9) reveals two sources of improvements in a farm's environmental performance. TE refers to the proportional decrease in inputs, while NAE relates to input combinations that have lower nutrient amounts. The values of these three efficiency measures are bounded between zero and one. The value of unity indicates full efficiency, whereas less than unity implies inefficiency.

As noted in the literature, the decomposition of NE and cost efficiency into a common TE component and allocative components (i.e., NAE and CAE) is particularly useful for analysing economic-environmental trade-offs (Lauwers, 2009; Van Meensel et al., 2010) and for policy implications (Coelli et al., 2007; Hoang and Coelli, 2011). Conceptually, an improvement in TE will yield higher cost and environmental efficiency levels. However, once the farms become technically efficient (i.e. no opportunity for TE increases) there exist a (negative) trade-off between cost and environmental allocative efficiency but the magnitude of trade-offs could vary as shown in Figure 1.

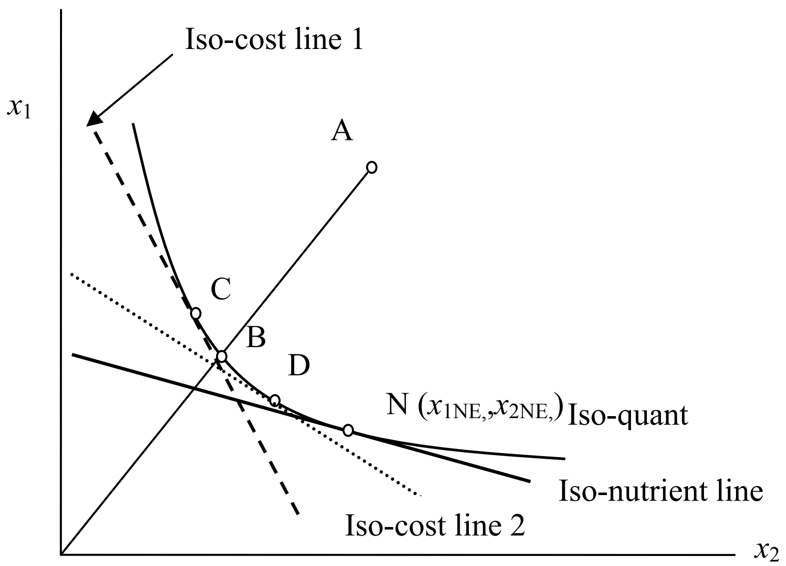

Figure 1: Trade-offs between cost and environmental efficiency

Figure 1 predicts an iso-nutrient line (a solid line) and two possible iso-cost lines (dot and dashed lines) for an observed farm (point A) given its iso-quant curve in a simple two inputs situation. Points B and N are technically and nutrient efficient points respectively. The two iso-cost lines are tangent to the isoquant curve at points C and D. Point A by moving to point B can improve TE. A move from points B to N represents an improvement in NAE. If the farm A faces the dashed iso-cost line, there is a clear negative trade-off between NEA and CAE: the farm will reduce costs by moving from points B to C but this move increases nutrient balance (because nutrient balance at point C is bigger that the balance at point B). If the farm A faces the dot iso-cost line, a move from points B to D reduces both costs and nutrient balance. However, a move from points D to N represents a trade-off between NAE and CAE. In this figure, the trade-off is much greater between points C and N than between points D and N.

Figure 1 also illustrates a useful case for possible policy intervention. Policies can affect the relative prices of inputs; hence the slope of the iso-cost line. In this example, a policy which changes the dashed to the dot iso-cost line is preferred because this policy could encourage farms to improve their environmental allocative efficiency and by doing this they can also improve cost allocative efficiency. Ideally, if the iso-cost line is identical to the iso-nutrient line, farms can be both environmental and cost efficient (Coelli et al., 2007; Hoang and Coelli, 2011).

As mentioned earlier, several empirical studies have estimated the NE in agricultural production (Coelli et al., 2007; Hoang and Alauddin, 2012; Hoang and Coelli, 2011; Nguyen et al., 2012; Reinhard and Thijssen, 2000; Van Meensel et al., 2010). These studies have focused on measuring efficiency levels and analysing economic-environmental trade-offs. These studies found high variation in nutrient efficiency across DMUs and a trade-off between cost efficiency and MB-based EE. None of these studies, however, investigated the determinants of variation. For policy analysis purposes, analysing factors that drive EE is important because such analysis could show directions for decision makers to find ways to improve environmental performance. Hence, we aim to fill this gap in the literature by utilising the two common analytical frameworks of analysing efficiency drivers. Details are discussed in the following section.

5. Analysing environmental efficiency variation

The most common approach to analysing efficiency determinants is a conventional two-stage DEA method. Efficiencies are estimated in the first stage and the estimated efficiencies are then regressed on explanatory variables in the second stage using ordinary/general linear least square or a censored (Tobit) model. Formally, econometric models in the second stage take the following form:

(11) $$\mathrm{NE}_{it} = \mathbf{z}_{it}\boldsymbol{\delta} + \boldsymbol{\varepsilon}_{it}$$

where NE_{it} refers to the values of nutrient-orientated EE of the i-th farm at period t-th, $\mathbf{z}_{it}$is a vector of explanatory variables, $\boldsymbol{\delta}$ is a vector of unknown coefficients to be estimated, and $\boldsymbol{\varepsilon}_{it}$ refers to the error terms. $\mathbf{z}_{it}$ can be interpreted as determinants of NE variation,

While this conventional two-stage approach is simple, it suffers from three main shortcomings. First, the data noise is included in DEA efficiency scores estimated in the first stage. Second, inconsistency exists because in the first stage efficiencies are assumed to be independently, identically distributed (iid) but in the second stage they are assumed to have a functional relationship with explanatory variables (Battese and Coelli, 1995). Third, statistical inferences regarding the significance of explanatory variables in explaining efficiency variation in the second stage is invalid due to a complicated, unknown serial correlation among the estimated efficiencies (Simar and Wilson, 2007).

Simar and Wilson (2007) propose two bootstrap algorithms to permit valid statistical inferences of the second-stage results. The first algorithm bootstraps the confidence intervals for improved statistical inferences. The second algorithm removes biases in efficiencies and bootstraps the confidence intervals of the coefficients of explanatory variables in a truncated regression model in which the bias-corrected inefficiencies are a dependent variable.

The second estimation method is conducted in a fully parametric SFA framework that estimates the parameters of the stochastic frontier and inefficiency models in a single stage. The stochastic frontier model estimates inefficiencies, while the inefficiency model identifies drivers of efficiency variation. This method can remove data noise from inefficiencies, but it may suffer from the problem of misspecification of functional forms. Similar stochastic models have been proposed for cross-sectional data (Kumbhakar et al., 1991) and panel data (Battese and Coelli, 1995). The present article proposes the use of this method to construct the stochastic nutrient frontier function (SNFF) for panel data sets, that is, to rewrite (4):

$$(12)\quad \mathrm{NC}_{it} = nc(\mathbf{q}_{it}, \mathbf{a}_{it}; \boldsymbol{\beta}).\exp\{V_{it} + U_{it}\}$$

where NC_{it} denotes the total amount of nutrients contained in inputs at the t-th observation for the i-th farm, $\mathbf{q}_{it}$ refers to outputs, $\mathbf{a}_{it}$ refers to the nutrient contents of inputs, and $\boldsymbol{\beta}$ is a vector of unknown parameters to be estimated. The V_{it}s are assumed to be iid $N(0,\sigma_v)$ random errors, independently distributed of the U_{it}s, which are non-negative random variables associated with nutrient inefficiency of production. The SNFF in (12) can be viewed as a counterpart of a cost frontier function. The nutrient frontier function attempts to minimise the total amount of nutrients used in production, whilst the cost frontier function minimises the total costs of production. The U_{it} in (12) defines how far above the nutrient frontier the farm operates, and it is assumed to have a relationship with explanatory variables **z** as in (11).

Note that both the truncated regression analysis and the inefficiency model impose a specific distribution form of the error terms (i.e., normality of the error term). One can argue that this imposition reduces the advantages of the non-parametric nature of DEA in the first stage. As there is a lack of theoretical background to choose between the two methods, we suggest researchers use both approaches in empirical studies, as we did, with the goal of providing more robust evaluations of the results.

6. An empirical study of rice production in Korea

Nguyen et al. (2012) used DEA to estimate the NE of paddy farms in the Gangwon province of South Korea between 2003 and 2007. This study estimated an average NE of 0.309, suggesting that these farms on average should be able to produce their current output with an input bundle that contains 69.1% less eutrophying power aggregated from N and P with the weights of 1:10. Their study also determined that NE varied significantly across farms. More importantly, a high correlation between cost efficiency and NE was observed, which led to the hypothesis that farms were facing a trade-off between economic efficiency and environmental performance during the period surveyed. In response to this study, we used the same dataset to analyse the determinants of NE variation and to revisit the hypothesis of trade-off between cost and nutrient efficiencies.

6.1 Data description and model specifications

Detailed descriptions of the dataset are in Nguyen et al. (2012). The dataset has 468 observations of 96 rice farms between 2003 and 2007. Outputs of net rice grain and straw were aggregated into a single measure of output using the Fisher quantity index with prices used as weights. The four groups of aggregated inputs include land (measured in ha), labour (measured in hours), fertilisers (a price-weighted Fisher quantity index aggregated from 19 different types of chemical and organic fertilisers), and other inputs (a price-weighted Fisher quantity index aggregated from 27 types of other inputs). N and P were the two main nutrients considered in this study. Scientific and experimental data of the polluting impacts of N and P on surrounding areas were not available. Instead of using one single set of weights (1N:10P) as in Nguyen et al. (2012), we followed the suggestions of Hoang and Coelli (2011) and used two other sets of weights (1N:1P and 1N:5P) to investigate how sensitive the results would be in relation to differing choices of N:P weights.

We estimated the SNFF and the inefficiency model in the single-stage SFA framework using Cobb-Douglas (CD) and translog functional forms for (11) using FRONTIER 4.0 package (Coelli, 1996b). CB and translog are two common functional forms used in empirical agricultural efficiency studies. In comparison with CD, translog is more flexible but has more parameters to estimate and hence, may give rise to econometric difficulties (Coelli et al., 2005). Given that the choice of functional forms can have impact on statistical inferences (Giannakas et al., 2003), we presented the results of both functions without drawing any definite conclusions with regards to which form is superior.

As in Reinhard and Thijssen (2000), labour and other inputs were considered as fixed inputs. Only land and fertilisers have nutrient contents. The CD and translog nutrient frontier functions (farm *i*-th and time period *t*-th subscripts were omitted for the sake of simplicity) are:

$$(13)\quad \ln nc = \beta_0 + \sum_{n=1}^{2} \beta_n \ln a_n + \kappa \ln q + \sum_{m=1}^{2} \varphi_m \ln f_m + u + v$$

$$(14)\quad \begin{aligned} \ln nc = \; & \beta_0 + \sum_{n=1}^{2} \beta_n \ln a_n + \kappa \ln q + \sum_{m=1}^{2} \varphi_m \ln f_m \\ & + 0.5\left(\sum_{n=1}^{2} \beta_{nn} \ln a_n^2 + \kappa_{11} \ln q^2 + \sum_{m=1}^{2} \varphi_{mm} \ln f_m^2 \right) \\ & + \beta_{12} \ln a_1 \ln a_2 + \sum_{n=1}^{2} \rho_n \ln a_n \ln q \\ & + \sum_{n=1}^{2} \sum_{m=1}^{2} \eta_{nm} \ln a_n \ln f_m + \sum_{m=1}^{2} \varsigma_m \ln q \ln f_m + \varphi_{12} \ln f_1 \ln f_2 + u + v \end{aligned}$$

where a_1 and a_2 refer to nutrient contents of land and fertilisers, f refers to fixed inputs (labour and others) and q refers to the output. After imposing linear homogeneity in nutrient contents, these nutrient frontier functions become:

$$(15)\quad \ln y = \beta_0 + \beta_1 w + \kappa \ln q + \sum_{m=1}^{2} \varphi_m \ln f_m + u + v$$

$$(16)\quad \begin{aligned} \ln y = \; & \beta_0 + \beta_1 w + \kappa q + \sum_{m=1}^{2} \varphi_m f_m + \beta_{12} w_{12} + 0.5\kappa_{11} \ln q^2 + \sum_{m=1}^{2} 0.5\varphi_{mm} f_m^2 \\ & + \rho_n wq + \sum_{m=1}^{2} \eta_{nm} w f_m + \sum_{m=1}^{2} \varsigma_m f_m q + \varphi_{12} f_1 f_2 + u + v \end{aligned}$$

where

(17) $\ln y = \ln nc - \ln a_2$,

(18) $w = \ln a_1 - \ln a_2$,

(19) $w_{12} = \ln a_1 \ln a_2 - 0.5(\ln a_1)^2 - 0.5(\ln a_2)^2$,

and u depends on explanatory variables as in (11).

Note that these frontier functions do not include trend terms. We exploratory estimated these frontiers with the trend terms, but the results did not suggest that the trend terms were significant (i.e., t-tests were not significant at either the 5% or 10% LOS, and log-likelihood tests could not reject the null hypotheses that favour the model without the trend terms).

We used the two-stage DEA method with the single-bootstrap algorithm of Simar and Wilson (2007). In the first stage of the DEA method, we used DEAP (Coelli, 1996a) to calculate NE under a constant return-to-scale (CRS) technology[2]. The reciprocal of NE is the dependent variable in the single-bootstrap truncated inefficiency model in (11) using the algorithm 1 of Simar and Wilson (2007) with 1000 replications in Matlab[3].

---

[2] Using a Cobb-Douglas production function, a chi-squared test confirmed that the production technology exhibits CRS (a test statistic of 0.644 < critical values at the 5% LOS = 3.84). The DEA efficiency results under the VRS specification were found greater than CRS (i.e., using a 1N:10P weights, the mean VRS's NE was 0.455). Detailed VRS results can be provided upon request to the authors.

[3] Our Matlab codes were modifications of the codes written by Zelenyuk and Zheka (2006).

Reinhard et al. (2002) present a conceptual scheme of a comprehensive model of farming in which all factors related to input and output qualities, production technologies, institutions, farmers' characteristics, and physical environments can be captured either in the first stage of efficiency estimation or in the second stage of efficiency variation analysis. Several socio-economic factors, such as the educational background, the gender, the employment status (full-time or part-time) of farm owners-managers, and the level of governmental subsidies[4] were found to be significant in explaining efficiency variation (see Bravo-Ureta et al. (2007) for a metadata analysis). However, these types of information were not available in the current data set. This lack of data may have caused biases in the results, which suggests that the reported results should be interpreted with caution.

The following six groups of explanatory variables were used: (1) total land area (measured in ha); (2) share of owned land out of the total land area (%); (3) fertiliser consumption intensity (kg/ha); (4) cost allocative efficiency; (5) three dummies (with ten year intervals) representing four age groups of the farm owner; and (6) nine dummies representing ten geographical regions. CAE, defined as the ratio of cost efficiency to technical efficiency in

[4] Subsidies for chemical fertilisers had been gradually reduced since 2003 and completely terminated in 2005. However, South Korea's government retained subsidies for organic fertilisers and other types of subsidies during the period surveyed (Kang and Kim, 2008). Information about these subsidies was not available for each farm. We tried to use a single dummy for chemical fertiliser subsidy to capture some of these policies, but the results did not suggest the significance of this dummy variable. We decided to report results of models without this dummy variable.

Nguyen et al. (2012), refers to the cheapest combination of inputs given various input combinations that produce same output levels. The base group for age dummies refers to the group of farms whose owners are 70 years old and above. The proportions of these four age groups (less than 50 years old, between 50 and 60 years, between 60 and 70 years, and 70 years old and above) are 18%, 23%, 40%, and 19%, respectively. Regional dummies are included to capture all other factors that may have affected NE variation, such as weather conditions, soil conditions, water availability, sunlight, etc., and the quality of all other inputs. Descriptive statistics of these explanatory variables are summarised in Table 1. These explanatory variables have often been used in empirical studies in the agricultural sector, with the exception of CAE.

Using the same data set, Nguyen et al. (2012) provide estimates of the magnitudes of trade-offs between cost efficiency and NE for each individual farms. This study reported that, on average, the movement from the cost efficient position (i.e. points C or D in Figure 1) to environmental efficient position (i.e. point N in Figure 1) would reduce nutrient use by about 64% but increase the cost by 259% whilst the opposite movement (i.e., NE position to CE position) would increase the nutrient consumption by 224% but reduce the costs by 57%. In this article, we are not interested in the magnitude of trade-offs but we wish to provide a statistical test for such a trade-off in the entire dataset. We hypothesise that NE may be affected by CAE.[5]

---

[5] We did not use the cost efficiency (CE) because as discussed earlier both CE and NE can be decomposed into the common TE component; hence including CE would not allow us investigate directly trade-off between NE and allocative cost efficiency.

Table 1: Descriptive statistics of explanatory variables

Variable	Mean	Std. Dev.	Min	Max
Total land area (ha)	1.376	1.459	0.212	9.392
The share of owned land in the total land area (%)	62.394	37.829	0	100
Fertiliser consumption intensity (kg/ha)	3159.186	2695.968	97.752	15153.28
Cost allocative efficiency	0.717	0.154	0.194	1.000

6.2 Nutrient efficiency

Table 2 provides descriptive statistics for technical efficiency (TE), nutrient allocative efficiency (NAE) and nutrient efficiency (NE) using three N:P settings calculated from the input-orientated DEA estimation under the CRS technology. The results exhibit high variation with respect to NE across farms. The results from the stochastic nutrient frontiers also indicate high variation in the NE across farms, as shown in Appendix 1[6].

Table 2: Mean DEA efficiency scores under three settings of N:P weights

		1N:1P		1N:5P		1N:10P	
	TE	NAE	NE	NAE	NE	NAE	NE
Geometric mean	0.753	0.642	0.484	0.665	0.500	0.349	0.263
Arithmetic mean*	0.771	0.660	0.503	0.682	0.519	0.396	0.309
Standard deviation	0.163	0.141	0.133	0.140	0.134	0.192	0.179
Min	0.283	0.087	0.057	0.094	0.062	0.077	0.055
Max	1.000	1.000	1.000	1.000	1.000	1.000	1.000
*: at arithmetic mean values, NE does not necessarily equal to TE x NAE							

The mean TE score of 0.753 suggests that 96 farms, on average, should be able to produce their current output quantities with 24.7% fewer inputs. When 1N:5P was used, the mean NAE score of 0.665 indicated that these farms could reduce the total N and P eutrophying

[6] DEA results allow the direct decomposition of NE into TE and NAE, whereas the results of SNFF do not allow such compositions. This is an advantage of the DEA framework.

power by 43.5% if they were to adjust the combination of nutrient-containing inputs (land and different types of fertilisers). The overall NE score of 0.5 implies that these farms should be able to produce the same output levels with inputs containing 50% less N and P eutrophying power.

NE scores changed to 0.484 and 0.263 in 1N:1P and 1N:10P scenarios respectively. In all three N:P settings, NE scores of less than 0.5 suggested that there were great opportunities for these farms to improve the efficiency of nutrient usage. Higher NE scores implied a less damaging eutrophying effect of aggregate N and P balances on the waterways. By improving NE, these farms could reduce potential eutrophication in water systems.

Table 3 reports the inter-temporal changes of TE, NAE and NE scores, and Figure 2 displays the movements of levels of outputs and three nutrient-containing inputs (land, chemical fertilisers, and organic fertilisers) between 2003 and 2007. The year 2005 was of special interest because subsidies for chemical fertilisers were halted during this year. Outputs were lower in 2006 and 2007 than they were in 2005 regardless of the expansion of land. As expected, the consumption of chemical and organic fertilisers was reduced after 2005; however, the consumption of organic fertilisers exhibited a much greater magnitude in reduction.

6.3 Sensitivity of NE to N:P weights

The tests in Appendix 2 confirmed that differences in DEA-based NE across the three settings were significant, implying that the NE scores are sensitive to N:P weights. Table 4 presents the estimates of the stochastic nutrient frontiers in (16), and the chi-squared tests prefer the translog models to the CD models in all three settings of N:P weights. As evidenced in Appendix 3, there are significant differences in NE scores between CD and translog forms in all three weight settings. Appendix 4 also shows that differences in NE scores between N:P weight settings were significant for the translog model and for the CD model. Surprisingly, as presented in the next section, the magnitude of relationships (i.e., the coefficient value) and the statistical inferences regarding the determinants of NE variation were not sensitive to the N:P weights. One possible explanation is the dominant affects of CAE on NE relative to other variables as shown in Appendices 5 and 6.

Table 3: Geometric annual mean values of efficiency measures

Year	ITE	INAE (1N:1P)	INAE (1N:5P)	INAE (1N:10P)	INE (1N:1P)	INE (1N:5P)	INE (1N:10P)
2003	0.764	0.734	0.755	0.528	0.561	0.577	0.404
2004	0.780	0.616	0.642	0.438	0.481	0.500	0.341
2005	0.740	0.614	0.638	0.259	0.454	0.472	0.191
2006	0.768	0.629	0.661	0.433	0.483	0.508	0.333
2007	0.715	0.624	0.633	0.200	0.446	0.452	0.143

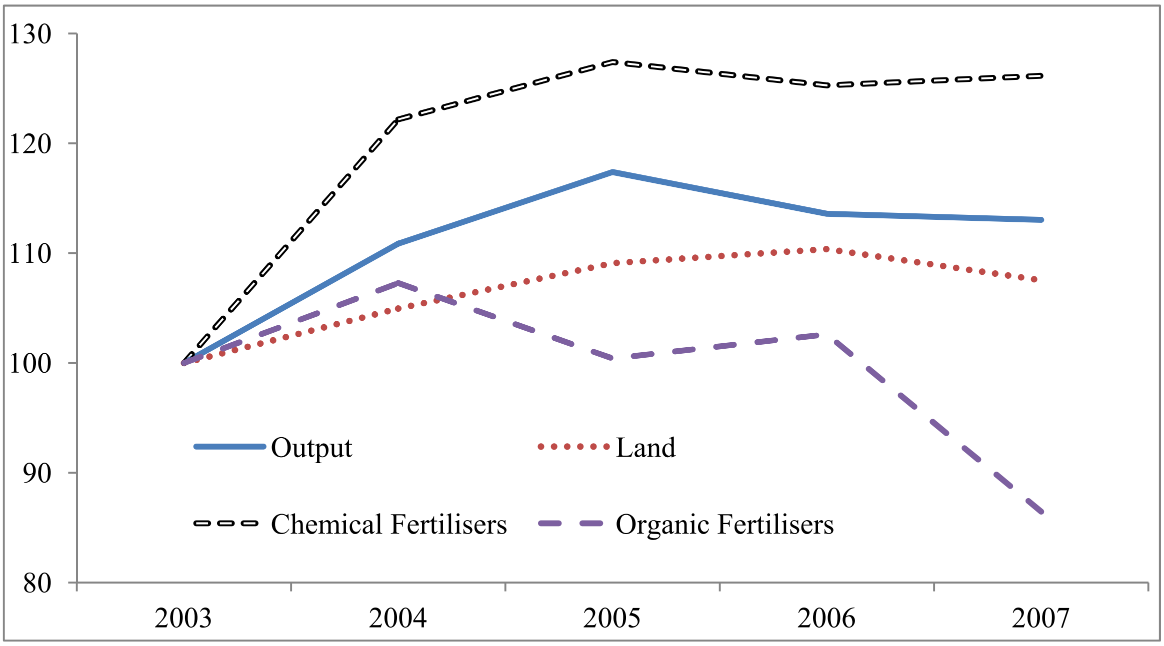

Figure 2: Indices of outputs, land area, chemical fertilisers and organic fertilisers

6.4 Determinants of nutrient efficiency variation

Table 4 summarises all statistical inferences from three estimations of equation (11): single-bootstrap truncated regression, CD inefficiency model, and translog inefficiency models. The dependent variable in the single-bootstrap truncated regression model was the reciprocal of the NE estimated by the input-orientated DEA framework, while the CD and translog inefficiency models had (1-NE) as a dependent variable. Hence, the magnitudes of the relationships (i.e., the absolute values of coefficients) in the three models were not directly comparable. Detailed results of these models are in Appendices 5 and 6.

The results in Table 5 show that the statistical inferences were highly consistent between three N:P weight settings in all three of the models. These results suggested that the determinants of NE are not sensitive to the choice of N:P weights in our empirical study. The CD and translog models reported similar signs of the coefficients for those variables that are statistically significant at the 5% or 10% LOS. However, minor differences existed between the truncated model and the CD and translog models.

Total land, the share of owned land, fertiliser consumption intensity, cost allocative efficiency and a dummy for region six were found to have significant relationships with nutrient inefficiencies at the 5% or 10% LOS. The directions of the relationship for the share of owned land, cost allocative efficiency and region 6 dummy were the same in all models. The signs for total land size and fertiliser consumption intensity were positive in CB and translog models and the 1N:10P version of the truncated regression model but the signs were negative in the truncated regression models using 1N:1P and 1N:5P settings. It is possible that these differences were caused by econometric issues and data noise. For example, bootstrapping may add noise to the estimates or these models may suffer from other unknown complicated econometric problems. These issues are beyond the focus of the present article but are worthy of further investigations in the future.

The nutrient inefficiencies were positively correlated with land size but negatively correlated with the share of owned land. These relationships suggested that those farms having larger land size and more rented land are less efficient in using nutrients. In South Korea, in comparison with part-time farmers, full-time farmers often cultivated on larger land areas with more rented land; hence, these two explanatory variables could be grouped together to represent the level of the commitment of the farms' owners and their families to the farm business. One can expect that those farmers with a stronger business commitment would be more technically efficient in using nutrient and non-nutrient inputs and that higher TE would yield better EE. Empirically, this observation was sensible in South Korea, given that TE scores were found to positively correlate with land (Kang and Kim, 2008).

CAE was found to be positively correlated with nutrient inefficiencies, implying that farms that are more cost allocative efficient tend to be less environmentally efficient; hence, there was a trade-off between cost and environmental efficiencies. As demonstrated in Nguyen et al. (2012), both nutrient and cost efficiency can be decomposed into technical efficiency and allocative terms (i.e., nutrient allocative and cost allocative). Thus a positive relationship between the nutrient inefficiency indicator and the CAE implied that farms that chose cheaper combinations of inputs generated greater balances of nutrients into the water system. Traditional economic theories suggest that, if the markets were free from distortions and captured well environmental pollution, the CAE could go hand-in-hand with environmental allocative efficiency. Obviously, the input markets that South Korean farmers have faced were highly distorted due to heavy subsidies from the governments (OECD 2008b). This finding posed an important policy implication that agricultural policies could be designed to affect the markets in a manner that allows farms to achieve more cost allocative efficiency and better environmental performance at the same time.

We could expect a positive relationship between fertiliser consumption intensity and nutrient inefficiencies. This expectation was found in CD and translog inefficiency models in all N:P settings and in the truncated regression model with the 1N:10P scenario. To reduce the nutrient balance that infiltrates the water system, farms should reduce the use of both chemical and organic fertilisers. Subsidies not only provide financial incentives for farmers to overuse fertilisers, but they also distort the markets of inputs in which farmers, without the proper understanding of agronomic knowledge, may use more fertilisers than other inputs.

Age dummies showed mixed results in the truncated regression, the CD and the translog inefficiency models, even though, in most of cases, these variables were not significant. The latter models (CB and translog inefficiency models) suggested that younger owners managed their farm to achieve better environmental performance than did owners aged between 60 and 69 years. Theoretically, this relationship is supported by the argument that older farmers tend to be more knowledgeable about past production technologies and less knowledgeable about the recent environmentally friendly production technologies (Weersink et al., 1990) and that younger farmers may be more open to adapt new technologies that are friendlier to the environment. This observation was sensible in South Korea, where young farmers were better educated and have had better access to different sources of information, including the Internet, to learn how to use nutrients more efficiently.

The empirical results also indicated that there was variation in the nutrient-orientated environmental performance of farms across different regions in the Kangwon province. The interpretation of these results, however, is not straightforward because there were many factors that these dummies were designed to capture (i.e., land quality, weather conditions, local market conditions, and so forth). If data about these variables are available, one could incorporate them into the model, thus allowing the results to deliver more practical interpretations.

7. Conclusions

There are two important components in empirical EE studies: the estimation of EE and the analysis of EE variation. The quality of the second component is critically determined by the reliability of the first component. The MB-based EE measures are argued to be particularly suitable for environmental analysis of farms. This article demonstrated that various econometric methods could be used to analyse the determinants of variation in the MB-based EE. Specifically, the bootstrap truncated two-stage DEA and stochastic nutrient frontier were proposed in this article. Moreover, economic-environmental trade-offs can also be statistically testified in these econometric models.

The empirical study in this article investigated variation in the MB-based EE across 96 farms in the Kangwon province of South Korea between 2003 and 2007. The land size of farms, fertiliser consumption intensity, cost allocative efficiency, and the share of rented land were found to have negative relationships with nutrient efficiency. These findings proffered two important implications. First, those farms with more business commitment tended to use more nutrients. Second, a negative relationship between cost-allocative and nutrient-allocative efficiency measures was identified, which suggests that farms were facing a trade-off between choosing cost-effective and environment-friendly combinations of inputs. Given high distortion of fertiliser markets due to government subsidies, these findings favoured the proposition that the government policies could be designed to change the behaviours of farms so that farms can achieve higher cost allocative efficiency and environmental efficiency simultaneously. The results also implied that removing fertiliser subsidies is not sufficient to discontinue the overuse of fertilisers. Other policy options should be considered to have stronger impacts on farms' behaviour. Examples of policies that need further analysis include policies that target reducing the balance of nutrients (for example imposing an environmental tax on the balance of nutrients) or educational or extension programs that provide farms with better on-farm nutrient management.

In this empirical study, the choice of N:P weights did not affect the statistical inferences even though they impacted the estimates of NE. The choice of econometric methods (either single-stage stochastic frontier or two-stage DEA), however, yielded differences in the statistical inferences. It is possible that these differences can be caused by a non-monotonic relationship between cost and environmental allocative efficiency, the absence of other important explanatory variables, and other unknown econometric problems. Hence, we suggested that researchers should take caution when deciding which method should be used. Undoubtedly, reporting the results from both methods is recommended.

Table 4: Stochastic nutrient frontiers

Variables	1N:1P				1N:5P				1N:10P			
	Cobb-Douglass		Translog		Cobb-Douglass		Translog		Cobb-Douglass		Translog	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Constant	0.181	0.002	-6.120	-0.111	0.747	2.932	1.901	-0.861	0.181	1.183	2.725	1.107
w	0.505	10.493	5.957	7.472	0.177	4.039	1.350	-2.928	0.505	2.077	-0.343	-0.682
Output	0.635	14.319	0.991	1.548	0.598	12.331	1.095	-1.893	0.635	11.459	0.836	1.198
Labour	0.190	4.159	-1.147	-1.431	0.244	5.025	-0.974	1.342	0.190	5.171	-0.663	-0.755
Others	0.039	3.230	0.274	1.272	0.048	3.187	0.249	-1.283	0.039	3.282	0.186	0.778
w12			1.688	8.197			0.683	-5.189			0.014	0.122
Output2			-0.007	-0.068			-0.099	0.953			-0.200	-1.525
Labour2			0.312	2.247			0.206	-1.520			0.137	0.808
Other2			-0.011	-1.481			-0.010	1.110			-0.008	-0.753
w*Output			0.122	1.374			0.009	-0.129			-0.003	-0.042
w*Labour			0.010	0.098			0.110	-1.357			0.122	1.282
w*Other			0.045	1.193			0.001	-0.030			-0.031	-0.829
Output*Labour			-0.212	-1.902			-0.130	1.210			-0.053	-0.391
Output*Other			0.048	1.405			0.015	-0.449			-0.013	-0.319
Labour*Other			-0.059	-1.679			-0.024	0.666			0.000	-0.001
sigma-squared	0.059	15.412	0.047	15.159	0.059	15.003	0.051	-15.588	0.059	15.185	0.076	13.971
gamma	0.531	0.043	0.523	0.037	1.000	16.248	1.000	-0.446	0.531	17.259	0.999	2.236
Log likelihood	-0.927		49.128		-1.853		34.098		-79.257		-60.950	
LR test	208.255		274.976		331.450		387.868		398.631		400.843	
Mean efficiency	4.850		5.671		4.280		5.403		6.131		7.538	
Chi-squared test*	100.1086688				71.9022006				36.613732			
* -2*[log(likelihoodH_0)-log(likelihoodH_1)]s follows a chi-squared distribution with 10 d.f. Critical chi-square=18.307 at the 5% LOS.												

Table 5: Statistical inferences in single-bootstrap truncated regression, Cobb-Douglas and translog inefficiency models

Variables	1N:1P			1N:5P			1N:10P		
	Single-bootstrap Truncated	Cobb-Douglass	Translog	Single-bootstrap Truncated	Cobb-Douglass	Translog	Single-bootstrap Truncated	Cobb-Douglass	Translog
Total land area	- , **	+ , *	+ , *	- , **	+ , *	+ , *	+	+ , *	+ , *
The share of owned land in total land	- , **	- , *	- , *	- , *	- , *	- , *	- , *	- , *	- , *
Fertiliser use per ha	- , *	+ , *	+ , *	- , *	+ , *	+ , *	+ , *	+ , *	+ , *
Cost allocative efficiency	+ , **	+ , *	+ , *	+ , **	+ , *	+ , *	+	+ , *	+ , *
Age (< 50 years)	- , **	+	+	- , **	+	+	- , **	+	+
Age (50-59 years)	-	+	+	-	-	+	-	+	-
Age (60-69 years)	- , *	+	+	- , **	-	+	- , *	-	-
Region 1	-	+	+	-	-	+	-	+	-
Region 2	-	+	+	-	-	+	-	+	+
Region 3	+	+	+	-	+	+	-	+	+
Region 4	-	+	+ , **	-	+	+	-	+ , *	+
Region 5	-	-	-	-	-	-	-	-	-
Region 6	+ , *	+ , *	+ , *	+ , *	+ , *	+ , *	-	+ , *	+ , *
Region 7	+	+ , **	+	-	+	+ , **	-	+	+
Region 8	-	-	-	+	-	-	-	-	-
Region 9	+	+	+	+	+	+	-	+	+

Note: +/- suggest a positive/negative relationship between nutrient inefficiency with explanatory variables; *: significant at the 5% LOS; **: significant at the 10% LOS. Detailed results are in appendices 5and 6.

Acknowledgements

This study was carried out as part of the International Research Training Group TERRECO (GRK 1565/1) funded by the German Research Foundation (DFG) at Bayreuth University, Germany and the Korean Research Foundation (KRF) at Kangwon National University, South Korea. We express our thanks to Bumsuk Seo for the data processing and Associate Professor Valentin Zelenyuk for giving us the Matlab codes.

References

Battese, G.E., Coelli, T.J., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical Economics 20, 325-332.

Callens, I., Tyteca, D., 1999. Towards indicators of sustainable development for firms: A productive efficiency perspective. Ecological Economics 28, 41-53.

Chung, Y.H., Fare, R., Grosskopf, S., 1997. Productivity and undesirable outputs: A directional distance function approach. Journal of Environmental Management 51, 229-240.

Coelli, T., 1996a. A guide to deap version 2.1: A data envelopment analysis (computer) program, The University of Queensland.

Coelli, T., Lauwers, L., Van Huylenbroeck, G., 2007. Environmental efficiency measurement and the materials balance condition. Journal of Productivity Analysis 28, 3-12.

Coelli, T.J., 1996b. Frontier version 4.1: A computer program for stochastic frontier production and cost function estimation. Center for Efficiency and Productivity Analysis, The University of Queensland.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battese, G., 2005. An introduction to efficiency and productivity analysis. Springer, New York.

Färe, R., Grosskopf, S., Pasurka, J.C.A., 2007. Environmental production functions and environmental directional distance functions. Energy 32, 1055-1066.

Färe, R., Grosskopf, S., Tyteca, D., 1996. An activity analysis model of the environmental performance of firms-application to fossil-fuel-fired electric utilities. Ecological Economics 18, 161-175.

Giannakas, K., Tran, K.C., Tzouvelekas, V., 2003. On the choice of functional form in stochastic frontier modeling. Empirical Economics 28, 75-100.

Greene, W., 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics 126, 269-303.

Hoang, V.-N., Alauddin, M., 2012. Input-orientated data envelopment analysis framework of measuring and decomposing economic, environmental and ecological efficiency: An application to OECD agriculture. Environmental and Resource Economics 51, 431-452.

Hoang, V.N., Coelli, T., 2011. Measurement of agricultural total factor productivity growth incorporating environmental factors: A nutrients balance approach. Journal of Environmental Economics and Management 62, 462–474.

Howarth, R., Anderson, D., Cloern, J., Elfring, C., Charles Hopkinson, Lapointe, B., Malone, T., Marcus, N., McGlathery, K., Sharpley, A., Walker, D., 2000. Nutrient pollution of coastal rivers, bays and seas. Issues in Ecology 7, 1-15.

Kang, H.-J., Kim, J.-H., 2008. Impact of direct income payments on productive efficiency of korean rice farms. Journal of Rural Development/Nongchon-Gyeongje 31, 1-22.

Kortelainen, M., 2008. Dynamic environmental performance analysis: A malmquist index approach. Ecological Economics 64, 701-715.

Kumbhakar, S.C., Ghosh, S., McGuckin, J.T., 1991. A generalized production frontier approach for estimating determinants of inefficiency in u.S. Dairy farms. Journal of Business and Economic Statistics 9, 279-286.

Kuosmanen, T., Kortelainen, M., 2005. Measuring eco-efficiency of production with data envelopment analysis. Journal of Industrial Ecology 9, 59-72.

Lauwers, L., Van Huylenbroeck, G., Rogiers, G., 1999. Technical, economic and environmental efficiency analysis of pig fattening farms, Poster presentation at the 9th European Congress of Agricultural Economists, Warchau, Polen.

Lauwers, L., 2009. Justifying the incorporation of the materials balance principle into frontier-based eco-efficiency models. Ecological Economics 68, 1605-1614.

Nguyen, T.T., Hoang, V.-N., Seo, B., 2012. Cost and environmental efficiency of rice farms in south korea. Agricultural Economics 43, 367-376.

Picazo-Tadeo, A.J., Beltrán-Esteve, M., Gómez-Limón, J.A., 2012. Assessing eco-efficiency with directional distance functions. European Journal of Operational Research 220, 798-809.

Picazo-Tadeo, A.J., Gómez-Limón, J.A., Reig-Martínez, E., 2011. Assessing farming eco-efficiency: A data envelopment analysis approach. Journal of Environmental Management 92, 1154-1164.

Reinhard, S., Lovell, C.A.K., Thijssen, G., 2002. Analysis of environmental efficiency variation. American Journal of Agricultural Economics 84, 1054-1065.

Reinhard, S., Thijssen, G., 2000. Nitrogen efficiency of dutch dairy farms: A shadow cost system approach. European Review of Agricultural Economics 27, 167-186.

Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136, 31-64.

Smith, V.H., Tilman, G.D., Nekola, J.C., 1999. Eutrophication: Impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. Environmental Pollution 100, 179-196.

Tyteca, D., 1996. On the Measurement of the Environmental Performance of Firms- A Literature Review and a Productive Efficiency Perspective. Journal of Environmental Management 46, 281-308.

Tyteca, D., 1999. Sustainability indicators at the firm level. Journal of Industrial Ecology 2, 61-77.

Van Meensel, J., Lauwers, L., Van Huylenbroeck, G., Van Passel, S., 2010. Comparing frontier methods for economic-environmental trade-off analysis. European Journal of Operational Research 207, 1027-1040.

Weersink, A., Turvey, C.G., Godah, A., 1990. Decomposition measures of technical efficiency for ontario dairy farms. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 38, 1023-1023.

Zelenyuk, V., Zheka, V., 2006. Corporate governance and firm's efficiency: The case of a transitional country, ukraine. Journal of Productivity Analysis 25, 143-157.

Appendix 1: NE results from SFA estimations

	Cobb-Douglas			Translog		
	1N:1P	1N:5P	1N:10P	1N:1P	1N:5P	1N:10P
Arithmetic Mean	0.206	0.234	0.163	0.176	0.185	0.133
Standard Deviation	0.076	0.144	0.145	0.090	0.145	0.142
Min	0.520	0.980	0.980	0.551	0.999	0.959
Max	0.077	0.083	0.051	0.035	0.036	0.022

Appendix 2: Paired t-tests on the differences of NE between N:P weight settings (DEA results)

	1N:1P & 1N:5P	1N:1P & 1N:10P	1N:5P & 1N:10P
H_0 (Mean 1 - Mean 2) =	0	0	0
Mean Difference	-0.016	0.194	0.210
Standard Error	0.012	0.197	0.193
p-value	0.000	0.000	0.000

Appendix 3: Paired t-tests on the differences of NE between Cobb-Douglas and translog

	1N:1P	1N:5P	1N:10P
H_0 (Mean 1 - Mean 2) =	0	0	0
Mean Difference =	-0.822	-1.122	-1.408
Standard Error	0.093	0.106	0.157
p-value	0.000	0.000	0.000

Appendix 4: Paired t-tests on the differences of NE between N:P weight settings (SFA results)

	Cobb-Douglas			Translog		
	1N:1P & 1N:5P	1N:1P & 1N:10P	1N:5P & 1N:10P	1N:1P & 1N:5P	1N:1P & 1N:10P	1N:5P & 1N:10P
H_0 (Mean 1 - Mean 2)	0	0	0	0	0	0
Mean Difference	0.570	-1.280	-1.850	0.269	-1.866	-2.135
Standard Error	0.034	0.102	0.076	0.044	0.146	2.453
p-value	1.000	0.000	0.000	1.000	0.000	0.000

Appendix 5: A single-bootstrap truncated regression inefficiency model

Variables	1N:1P			1N:5P			1N:10P		
	Coeff.	Low95%	Up95%	Coeff.	Low95%	Up95%	Coeff.	Low95%	Up95%
Constant term	-0.767	-4.149	2.921	0.041	-2.087	2.278	1.250	-2.969	5.633
Total land area	-0.406	-0.811	0.119	-0.333	-0.682	0.086	0.149	-0.387	0.657
The share of owned land in the total land area	-0.015	-0.029	-0.001	-0.013	-0.025	-0.001	-0.018	-0.036	-0.001
Fertiliser use per ha	-0.0003	-0.001	-0.0001	-0.0003	-0.0005	-0.0001	0.001	0.001	0.001
Cost allocative efficiency	3.218	-0.356	6.923	2.315	-0.662	4.997	-1.465	-6.323	2.956
Age (< 50 years)	-1.535	-3.216	0.240	-1.263	-2.669	0.257	-1.985	-4.091	0.063
Age (50 and 59 years)	-1.122	-2.683	0.361	-0.916	-2.131	0.471	-0.982	-2.918	1.133
Age (60 and 69 years)	-1.673	-3.139	-0.221	-1.470	-2.709	-0.120	-2.026	-3.718	-0.326
Region 1	-0.842	-2.702	1.330	-0.824	-2.342	1.096	-0.803	-3.486	1.772
Region 2	-0.015	-0.808	0.952	-0.426	-2.084	1.625	0.773	-2.003	3.600
Region 3	0.339	-1.374	2.003	-0.028	-1.218	1.342	-0.161	-2.167	1.880
Region 4	-0.653	-2.296	1.125	-0.642	-2.022	1.131	-0.459	-2.804	1.842
Region 5	-0.056	-1.820	1.862	-0.315	-1.760	1.313	-1.152	-3.709	1.481
Region 6	3.591	1.400	5.629	3.046	1.426	4.748	2.107	-0.769	5.149
Region 7	0.059	-1.752	2.183	-0.015	-0.949	0.790	1.313	-1.428	4.165
Region 8	-0.014	-0.844	0.878	0.207	-1.211	1.981	-1.223	-3.739	1.581
Region 9	0.506	-1.687	3.323	0.503	-1.346	2.748	0.871	-2.028	4.101

Appendix 6: Single-stage nutrient inefficiency models

Variables	Cobb-Douglas						Translog					
	1N:1P		1N:5P		1N:10P		1N:1P		1N:5P		1N:10P	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Constant term	0.820	0.010	0.447	1.579	0.522	1.948	0.674	0.010	0.414	1.365	0.488	0.895
Total land area	0.097	6.092	0.084	5.016	0.078	3.982	0.193	6.092	0.195	7.414	0.195	5.957
The share of owned land in total land	-0.001	-2.993	-0.001	-3.082	-0.001	-3.145	-0.001	-2.993	-0.001	-2.255	-0.001	-2.409
Fertiliser use per ha	0.0001	13.346	0.000	17.747	0.0001	21.942	0.000	13.346	0.0001	21.776	0.0001	19.250
Cost allocative efficiency	0.430	3.950	0.554	4.631	0.735	5.501	0.468	3.950	0.542	4.896	0.668	4.739
Age (< 50 years)	0.054	1.396	0.051	1.292	0.026	0.498	0.029	1.396	0.022	0.578	0.011	0.229
Age (50-59 years)	0.049	1.351	0.015	0.414	-0.015	-0.341	0.054	1.351	0.007	0.189	-0.005	-0.118
Age (60-69 years)	0.031	0.964	0.004	0.131	-0.011	-0.265	0.012	0.964	-0.025	-0.821	-0.018	-0.434
Region 1	0.023	0.454	0.010	0.202	-0.021	-0.360	0.029	0.454	0.014	0.293	-0.012	-0.184
Region 2	0.031	0.582	0.031	0.593	-0.001	-0.018	0.038	0.582	0.051	0.976	0.024	0.334
Region 3	0.072	1.473	0.073	1.513	0.040	0.810	0.070	1.473	0.065	1.339	0.038	0.609
Region 4	0.057	1.180	0.084	1.789	0.071	1.479	0.084	1.180	0.100	2.093	0.072	1.207
Region 5	-0.034	-0.668	-0.027	-0.540	-0.040	-0.673	-0.027	-0.668	-0.031	-0.621	-0.034	-0.490
Region 6	0.156	2.856	0.158	3.071	0.159	2.674	0.109	2.856	0.132	2.513	0.159	2.148
Region 7	0.110	1.942	0.093	1.603	0.065	0.944	0.085	1.942	0.067	1.171	0.063	0.824
Region 8	-0.034	-0.680	-0.016	-0.324	-0.026	-0.403	-0.043	-0.680	-0.032	-0.625	-0.034	-0.487
Region 9	0.066	1.163	0.066	1.136	0.039	0.532	0.068	1.163	0.047	0.801	0.027	0.403