# ASSESSMENT OF EFFICIENCY AND IMPACT OF SPECIFIC FACTORS ON WHEAT CULTIVATION IN UKRAINIAN AGRICULTURAL ENTERPRISES

**ROMAN SLASTON** and **KARIN LARSÉN**
Department of Structural Development of Farms and Rural Areas, Leibniz Institute of Agricultural Development in Eastern and Central Europe (IAMO), Germany
romasl@ukr.net, larsen@iamo.de

**Paper prepared for presentation at the 114th EAAE Seminar**
**'Structural Change in Agriculture', Berlin, Germany, April 15 - 16, 2010**

*Copyright 2010 by author(s). All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.*

## Abstract

Wheat is one of the most important cereals produced in Ukraine. This study aims to provide an overall picture of the efficiency in wheat production in three Ukrainian regions - Kiev, Poltava and Cherkasy - using data from 2006-08. This is done on the one hand by estimating different types of inefficiencies, using DEA, and to identify their determinants, and on the other hand by analyzing whether production factors are over- or underused. The findings suggest, among other things, that legal form often has an impact on efficiency and that variable production factors was underused during the analyzed time period.

**Keywords:** Efficiency, Ukraine, Wheat cultivation, Data Envelopment Analysis, Bootstrap, Tobit regression.

## 1. Introduction

Production of cereals constitutes a major part of Ukrainian agriculture. It corresponds to nearby 21 % of the gross production of the agrarian sector of the national economy. 48 % of all produced grain consists of wheat which is cultivated in almost every agricultural enterprise in Ukraine.[1]

After disintegration of Soviet Union, recession of volumes and production efficiency fall was observed in the agrarian sector of Ukraine as well as in the whole economy as result of inability to quickly adapt to market economy conditions and disagreements in a state policy. Not until year 2000, basic development indicators for agriculture started to display certain growth (primarily for production of grain and olive cultures). The subsequent development and favorable weather conditions contributed to reach a production level in the same size as that of year 1990 and to harvest a record yield of grain crops – 53.3 million tons, including wheat – 25.9 million tons at average productivity of 3.67 tons per hectare in 2008. Comparing to previous year which was not so favorable for crop production there was a growth of cereals production corresponding to 82%. This growth led to an immediate fall in cereal prices. As a result, some firms experienced losses from their grain cultivation despite the high land productivity.

Productivity of cereal production is sometimes evaluated by comparing yield per hectare with the average regional level. This is however a partial productivity measure which does not indicate total production efficiency. The reason is that wheat cultivation, just like the production of most

[1] Source: data from State Statistics Committee of Ukraine for 2008.

other goods, involves several different production factors and the efficiency of their use may differ substantially from farm to farm. In order to obtain a more accurate measure of efficiency - that considers multiple production factors - Data Envelopment Analysis (DEA) is applied in this study. Econometric tools will further be utilized in order to analyze the influence of various factors on efficiency.

The aim of this study is to provide an overall picture of efficiency in wheat production in three Ukrainian regions with relatively similar natural conditions: Kiev, Poltava and Cherkasy. This is on the one hand done by estimating efficiency in wheat production for three years (2006-08) and identifying the influence of various factors, such as enterprise size, specialization and legal form, on efficiency, and on the other hand by analyzing whether the different production factors are over- or underused. We furthermore discuss the impact for efficiency studies of using data on inputs and outputs in value terms rather than physical quantities and apply an alternative DEA cost efficiency estimator suggested by Tone (2002).

Analysis of the influence of factors such as size of farm, legal form, intensity of production, legal form, specialization and ratio of factors on farm efficiency was made by Lissitsa and Odening (2005). They utilized data of Ukrainian agricultural enterprises in a transition period (1990-1999) and found a positive influence of farm size and specialization in crop growing on technical efficiency.

As mentioned above, this study also aims to emphasize the importance of correct model specification and choice of input set which are used for producing outputs. When calculating the DEA efficiency scores, we utilize information about inputs in value terms rather than in physical quantities. This is often done in efficiency studies applied to agricultural production (e.g. Singh et al, 2000) as information about production factors often are more easily available in form of expenditures. However, in opposite to what is conventionally the case, we will not evaluate efficiencies with traditional DEA models under the assumption of constant input prices. The implications of using value-based DEA instead of physical quantities have recently been given attention in the literature (e.g. Cross and Färe, 2008; Tone, 2002). As illustrated by Tone (2002), in the case when firms face different prices there are some shortcomings of the traditional DEA cost efficiency measures. More specifically, it can be shown that in the case of two firms that uses the same amount of inputs but one facing input prices that are twice as high, the firms will have the same level of cost efficiency. This may seem a bit counterintuitive and one could expect that the firm that pays “too high” prices should have a lower cost efficiency score. Tone (2002) argues that this is a shortcoming of traditional DEA models and suggests new models which are slight modifications of the standard DEA models. These modified models are applied in this study.

The remainder of this paper is organized as follow. Section 2 mainly describe the new DEA models for evaluation new technical, allocative and cost efficiency suggested by Tone (2002). Section 3 discusses the data set and the definitions of variables used in this study. Empirical results are presented and discussed in section 4 followed by discussion and conclusions.

## 2. Methodology

Since the original idea of efficiency measurement was suggested by Farrell in 1957, the two main methods that have been used are data envelopment analysis (DEA) and stochastic frontier analysis (SFA). They involve mathematical programming and parametric (econometric) methods respectively. Two advantages of SFA over DEA are that it accounts for noise and facilitates conventional statistical tests of hypotheses. DEA also has a few advantages over SFA. For example, a distributional form of inefficiency effects or functional form for the production (cost) function does not have to be assumed. DEA also facilitates decomposition of inefficiency in an easy way. In this study we apply DEA because of the advantages mentioned above. To justify our methodological choice, obvious outliers in the dataset were excluded.

### 2.1 Principals of DEA efficiency measurements and the modified DEA model suggested by Tone (2002)

Farrell (1957) proposed that the efficiency of a firm consists of two components: technical, reflecting a firm’s ability to obtain maximal output from a given set of inputs, and allocative efficiency, reflecting a firm’s ability to use the inputs in optimal proportions given their prices and the production technology. These two measures can be combined to provide a measure of overall or cost efficiency.[2]

DEA is a linear programming approach which constructs a non-parametric piece-wise-linear convex isoquant consisting of the best performing firms in the sample. This method was developed and refined by a number of authors, including Charnes et al. (1978), Banker et al.

[2] Here we use terms “allocative” and “cost” efficiency with respect to resent literature. Originally, Farrell used terms “price” and “overall” efficiency, respectively.

(1984) and Färe et al. (1985; 1994). DEA models can have either input or output orientation. A description of the DEA methodology and the formulation of the general optimization problem can be found in a number of text books (e.g. Coelli et al., 2005).

For many types of production, including agricultural production, it is common to have information about output and utilized production factors in value terms (i.e. revenues and expenditures) rather than in physical quantities. As a result, it is not unusual that authors use (deflated) value terms instead of quantities when obtaining technical efficiency scores. In fact, the two models will produce identical efficiency scores when firms face identical prices. Recently, the difference between value-based and quantity based DEA models have been given attention in the literature (e.g. Cross and Färe, 2008; Tone, 2002). Tone (2002) identified some shortcomings of the original formulations of the DEA models for calculating cost efficiency when firms face different prices as it does not consider the possibility that a firm could pay lower input prices. As a results, two firms that use the same amount of inputs and produce the same amount of output but one of them pays twice as high price for the inputs will have the same economic efficiency. In order to resolve this shortcoming, the author suggested new DEA models which use value terms.

As we have inputs in value terms and unequal input prices across firms, we represent input-oriented DEA models developed by Tone (2002). He defined DEA linear program for estimation technical efficiency as follows:

$$\bar{\theta}_c^* = \min_{\bar{\theta},\lambda} \bar{\theta} \tag{1}$$

subject to

$$\bar{\theta}\bar{x}_o \geq \bar{X}\lambda \tag{2}$$

$$y_o \leq Y\lambda \tag{3}$$

$$\lambda \geq 0 \tag{4}$$

where $\bar{\theta}$ is a scalar, λ is a N×1 vector of constants, $\bar{x}_o$ and $y_o$ are observed input and output, respectively, $\bar{X}$ is N×M matrice of inputs, $Y$ is N×S matrice of output and N, M, S are numbers of firms, inputs and outputs in the sample. All firms belong to cost-based production possibility set $P_c$ :

$$P_c = \{(\bar{x}, y) \mid \bar{x} \geq \bar{X}\lambda, y \leq Y\lambda, \lambda \geq 0\}, \tag{5}$$

where $\overline{X} = (\overline{x}, ..., \overline{x}_n)$ with $\overline{x}_j = (c_{1j}x_{1j}, ..., c_{mj}x_{mj})^T$ and unit input cost $C = (c_1, ..., c_n)$. Here, we assume that the matrices *X* and *C* are nonnegative, and all inputs are associated with a cost. We further assume that the elements of $\overline{x}_{ij} = (c_{ij}x_{ij})(\forall(i, j))$ are denominated in homogeneous units, such as dollars, so that adding up the elements of $x_{ij}$ has a meaning. The value of $\overline{\theta}_c^*$ obtained will be the technical efficiency score for the i-th firm. It will take a value between 0 and 1, with a value of 1 indicating a point on the frontier and hence a technically efficient firm. The linear programming problem must be solved N times, once for each firm in the sample. A value of $\overline{\theta}_c^*$ is then obtained for each firm.

The linear program (LP) presented in (1)-(4) calculates efficiency under the assumption of constant return to scale (CRS). To obtain efficiency under the assumption of variable return to scale (VRS) the LP in (6)-(10) can be extended with constraint $e\lambda = 1$. The LP will then be as follows:

$$\overline{\theta}_v^* = \min_{\overline{\theta}, \lambda} \overline{\theta} \qquad (6)$$

subject to $$\overline{\theta}\overline{x}_o \geq \overline{X}\lambda \qquad (7)$$

$$y_o \leq Y\lambda \qquad (8)$$

$$e\lambda = 1 \qquad (9)$$

$$\lambda \geq 0 \qquad (10)$$

where $\overline{\theta}_v^*$ is the technical efficiency scores under VRS. Note that efficiency under variable return to scale can be no less than under constant return to scale, thus $\overline{\theta}_c^* \leq \overline{\theta}_v^*$ (for more details see Cooper et al. (2007)). Now we can decompose CRS technical efficiency ($TE_{crs}$) in to components: VRS technical efficiency ($TE_{vrs}$) and scale efficiency ($SE$). It can be expressed as:

$$TE_{crs} = TE_{vrs} \times SE. \qquad (11)$$

After evaluations of $\overline{\theta}_v^*$ and $\overline{\theta}_c^*$ scale efficiency can be obtained as:

$$SE = \frac{\overline{\theta}_c^*}{\overline{\theta}_v^*}. \quad (12)$$

One shortcoming of this measure of scale efficiency is that the value does not indicate whether the firm is operating in an area of increasing or decreasing return to scale. This latter issue can be determined by running an additional DEA problem with non-increasing (non-decreasing) return to scale imposed (e.g., Coelli et al 2005).

### 2.3 Cost and allocative DEA models

Traditional linear programs were used by many authors for determining cost and allocative efficiencies. However, it is only meaningful when input prices are constant across firms. If prices vary from firm to firm in the sample, traditional scheme lead to shortcomings and irrationality of the cost and allocative efficiencies.

Tone (2002) suggested that these shortcomings are caused by the structure of the supposed production possibility set $P$:

$$P = \{(x, y) \mid x \geq X\lambda, y \leq Y\lambda, \lambda \geq 0\} \quad (13)$$

$P$ is defined only by using technical factors $X = (x_1,...,x_n) \in R^{m\times n}$ and $Y = (y_1,...,y_n) \in R^{s\times n}$, and does not consider the unit input cost $C = (c_1,...,c_n)$.

Tone defined another cost-based production possibility set $P_c$ as:

$$P_c = \{(\overline{x}, y) \mid \overline{x} \geq \overline{X}\lambda, y \leq Y\lambda, \lambda \geq 0\}, \quad (14)$$

where $\overline{X} = (\overline{x},..., \overline{x}_n)$ with $\overline{x}_j = (c_{1j}x_{1j},...,c_{mj}x_{mj})^T$ under the assumption that the matrices X and C are nonnegative, and all inputs are associated with cost. Also he assumed that the elements of $\overline{x}_{ij} = (c_{ij}x_{ij})(\forall(i, j)$ are denominated in homogeneous units, for example dollars, so that adding up the elements of $x_{ij}$ has a meaning. Though we only assume here the convexity for the sets $P$ and $P_c$ as defined respectively by (13) and (14).

With production possibility set $P_c$ cost efficiency under VRS $\overline{\gamma}_v^*$ is defined as

$$\overline{\gamma}^* = e\overline{x}_o^* / e\overline{x}_o, \quad (15)$$

where $e \in R^m$ is a row vector with all elements being equal to 1, and $\bar{x}_o^*$ is the optimal solution of the LP given below:

$$\min e\bar{x} \tag{16}$$

subject to

$$\bar{x}_o \geq \bar{X}\lambda \tag{17}$$

$$y_o \leq Y\lambda \tag{18}$$

$$e\lambda = 1 \tag{19}$$

$$\lambda \geq 0. \tag{20}$$

After obtaining of technical ($\bar{\theta}^*$) and cost ($\bar{\gamma}^*$) efficiencies, allocative efficiency $\bar{\alpha}^*$ is then defined as the ratio of $\bar{\gamma}^*$ to $\bar{\theta}^*$, ie,

$$\bar{\alpha}^* = \bar{\gamma}^* / \bar{\theta}^*. \tag{21}$$

All efficiency measures $\bar{\theta}^*$, $\bar{\gamma}^*$ and $\bar{\alpha}^*$ are units invariant as long as $\bar{X}$ has a common unit of cost, e.g. Dollar, Euro or Pound. All of these measures are bounded by 0 and 1. The model in (16)-(20) searchs for the optimal input mix $\bar{x}^*$ for producing $y_o$ (or more). More concretely, the optimal mix is described as:

$$\bar{x}_i^* = \sum_{j=1}^{n} c_{ij} x_{ij} \lambda_j^* \qquad (i = 1,...,m) \tag{22}$$

Thus the optimal input mix for producing a given output $y_o$ can be found independently of the current unit cost $c_o$ of $DMU_o$. This is the fundamental differences between traditional model based on production possibility set $P$ in (13) and the model developed by Tone (2002). The traditional one cannot recognize the existence of other cheaper mix of inputs.

As the worldwide globalization of production has become a current trend, we should be able to find out the optimal input mix or at least to notify the existence of cheaper ones through the cost efficiency estimation.

### 2.4 Shadow prices and implicit value shares

Although the envelopment form of DEA model presented in (6)-(10) has good calculation properties and provides valuable information, this form does not provide direct information about shadow prices used for efficiency estimation. Shadow prices can be used for calculation of implicit value shares. These shares contain valuable information as they can be compared with actual cost shares and thereby indicate potential over or underuse of inputs.[3] If the implicit value share of an input is higher than the actual share, this suggests that the input was underused. If the opposite case is true, then the input was overused. In the case that the actual and shadow shares coincide, this implies that the input was used in optimal proportions. To obtain these shadow prices and shares, the dual form of linear program has to be solved.[4] A number of papers, including Charnes et al (1978), introduce the dual form by first presenting the more intuitive ratio form. In the ratio form one seeks to obtain an efficiency measure defined as the ratio of a weighted sum of all outputs over a weighted sum of all inputs. The optimal weights are obtained by solving the mathematical programming problem:

$$\max_{u,v}(u' y_i / v' x_i), \tag{23}$$

subject to $$u' y_j / v' x_j \leq 1, \tag{24}$$

$$u, v \geq 0, \tag{25}$$

where $y_i$ and $x_i$ are row vectors of outputs and inputs, respectively, for the *i*-th firm, $u'$ and $v'$ are row vectors of weights for outputs and inputs, respectively.

This involves finding values for $u$ and $v$, such that the efficiency measure for the *i*-th firm is maximized, subject to the constraints that all efficiency measures must be less than or equal to one. One problem with this particular ratio formulation is that it has an infinite number of solutions (see Coelli et al (1998)). To avoid this, one can impose the constraint $v' x_i = 1$, which provides:

$$\max_{u,v} \mu' y_i, \tag{26}$$

subject to $$v' x_i = 1, \tag{27}$$

[3] Coelli and Prasada Rao (2001) discuss implicit value shares in Malmquist TFP index numbers.
[4] Coelli and Prasada Rao (2001) mention however that the dual weights can be obtained from the simplex tablau of the enveleopment, or primal form of the DEA model.

$$\mu' y_j - \nu' x_j \leq 0, \quad (28)$$

$$\mu, \nu \geq 0, \quad (29)$$

where the change of notation from $u$ and $v$ to $\mu$ and $\nu$ is used to stress that this is a different linear programming problem. The form in equations (26)-(29) is generally known as the dual form of the DEA linear programming problem. The elements in the vectors $\mu$ and $\nu$ may be interpreted as normalized shadow prices. Thus, $\mu' y_i$ and $\nu' x_i$ are implicit value shares.

**3. Data**

The utilized data consists of 2110 firm-year observations from the regions Kiev, Poltava and Cherkasy covering the years 2006-2008. The regions Kiev, Poltava and Cherkasy were chosen as they are located geographically close to each other and have similar production conditions. In 2008, these regions produced close to 20 percent of total production of grain in Ukraine, and wheat production is widespread in analyzed area. The data set, provided by the association "Ukrainian Club of Agrarian Business", consists of farm level data and contains information about farming activities in each of the three years 2006, 2007 and 2008.

In 2007, very bad weather conditions caused a low level of productivity while year 2008 was very beneficial for cereals growing. Therefore, we expect to find differences of efficiency levels through this period. The three years was analyzed as one sample. Thus, if there is some relation between efficiency and specific factor (such as legal form) we assume that this will be observed independently of chosen period (one year or three years). To achieve comparability of data from different years, suitable price indices were used. These were obtained from the webpage of State Statistical Committee of Ukraine.

The original sample consisted of 3574 agricultural enterprises. A majority of the Ukrainian agricultural enterprises operate on rented land and we excluded those farms that have some share of owned land. In fact, agricultural land is not the subject of the trade in Ukraine and agricultural firms cannot buy it (it is forbidden by law)[5]. Thus, mainly public/state enterprises have non-rented (owned) land. This allowed us to measure the land input in monetary units and to apply the new DEA cost efficiency suggested by Tone (2002). After excluding these and cleaning for missing and inconsistent observations, 2110 farms remained. Of these, 613 were left in 2006, 708

[5] The Land Code of Ukraine of 25.10.2001 № 2768-III

in 2007 and 789 in 2008. Some descriptive statistics of the sample farms according to region can be found in Table 1.

***** Table 1 here *****

It can be noted that there are differences through regions and years. The region of Poltava has, on average, bigger farms compared to the other two regions (in terms of land size). Land productivity was higher in Cherkassy trough all years. This can be explained by a slightly better soil quality with higher share of humus in this region. The share of arable land used for wheat production was larger in the region of Kiev.

The utilized data set consists of detailed information about utilized inputs for the different crops, and this study will focus on the production of wheat. In the efficiency calculations, one output and four inputs is used. The output is the total quantity of produced wheat (in metric tons). The input variables represent the input usage attributable to the farms wheat production. All inputs expressed in value terms (UAH)[6]. The inputs considered are:

- labor: basic and additional payments to the workers, servants, managers directly engaged in technological process of wheat production;

- variable inputs: all material costs related to wheat production such as seed, fertilizers, pesticides, fungicides, herbicides and other variable inputs;

- machinery: fuel, depreciation of machines and services of other companies used for wheat producing;

- "other" costs: all costs which was not included in previous expenses like land rent, fixed agricultural tax, costs for missions, charges to social funds and other costs.

Thus, land was not included in the input set directly, but was implicitly taken into account in other costs as rent. As cost efficiency is of particular interest of our study, using land (measured in hectares) as input makes evaluation of it impossible. This is why all inputs were measured in value terms.

Descriptive statistics of the analyzed variables can be found in Table 2.

*** **Table 2 here *****

As we can see, a very wide range of firms was analyzed. In average farms have about 2.5 thousand hectares of agricultural land, but it varies from 23 to 70 thousand hectares. The average

[6] Ukrainian Hryvnia (UAH) – official currency of Ukraine. 1000 UAH = appr. 85 Euro (August, 2009)

share of land used for wheat growing is 23 %, what can be explained from the point of expedience of crop rotations. Specialization in crop growing is very high because the unprofitability of livestock production in recent years compelled many enterprises to sell their animals and to concentrate on cash crop production.

Beside these variables, information about legal forms of the enterprises was utilized in the analysis. The legal forms considered were limited liability companies, private enterprises, farms, cooperatives, and other forms. In Ukraine, the most common legal form for agricultural firms is the limited liability company. Number of firms and share of each type of legal form in each year are presented in Table 3.

*** **Table 3 here** ***

Table 4 displays some farm indicators (size, average wheat yield per ha and share of land used for wheat) by legal form and year.

*** **Table 4 here** ***

## 4. Results

The results of the DEA efficiency estimation are presented in Table 5. This table consists of average technical, cost, allocative and scale efficiency for the whole sample as well as for groups stratified by region and legal form. Technical efficiency scores are presented for both the assumption of constant returns to scale (CRS) and variable returns to scale (VRS). As efficiency scores usually are obtained for a sample of the population, which also is the case here, they are expected to be upward biased because of the so-called sample size bias (see for example Staat, 2002, for a discussion about this), we also report bias-corrected CRS and VRS efficiency scores obtained using the homogenous bootstrap suggested by Simar and Wilson (1998; 2000) obtained using FEAR (Wilson, 2008). Their procedure also facilitates calculation of confidence intervals. Average upper and lower level of 90%-confidence intervals are reported in Table 5. The calculation of confidence intervals and bias-corrected efficiency scores was done to justify deterministic nature of DEA efficiency scores and to provide us with more information for the conclusions.

*** **Table 5 here** ***

The average CRS technical efficiency score for our sample is 0.57 suggesting that there is a quite large heterogeneity among the farms with large potential room for efficiency improvements for many of them. It is however important to keep in mind that absolute level of the efficiency score is affected by factors such as the aggregation level of inputs and outputs (i.e. the number of inputs and outputs) and the total number of observations. In the region of Poltava, technical efficiency scores are higher on average than for the other two regions. Bias-corrected scores and confidence intervals of technical efficiency support significance of differences between the regions. Thus agricultural enterprises in the region of Poltava were more technically efficient on average than enterprises located in Cherkasy and Kiev. Enterprises in the Cherkasy region had on average lower efficiency than those located in Poltava but higher than those in Kiev, and the lowest average technical efficiency scores was observed for the enterprises in Kiev. The same situation is observed for cost efficiency which was 0.56 for the most productive region of Poltava. Allocative and scale efficiencies were quite high and almost similar for all regions. In relation to legal forms, the results suggest that technical efficiency of cooperatives is higher than for the other legal forms. For other legal forms, the results of bootstrapping do not indicate significant differences as their average confidence intervals are close to each other.

An interesting question is now: why are cooperatives more efficient than the other legal forms? Lissitsa and Odening (2005) found no influence of legal form on efficiency. It is however important to point out that they where looking at the transition period, and the term “cooperative” now has a slightly different meaning: it is not only legal form but also a different organizational form. Members of cooperatives are workers and owners at the same time. The profit is divided between the members according to their labor contribution in the production. As they work for themselves, each worker is therefore highly interested in productivity and profitability of their business.

As the efficiency scores were obtained as weak Farrell’s efficiency, it is advisable to solve phase II of linear programming problem (a detailed description can be found in Cooper et al. (2007)) which calculate slacks. The results of calculation of slacks according to legal form are demonstrated below.

*** **Table 6 here *****

The results presented in Table 6 suggest that labor is the most overused input in slacks sense, especially in state enterprises. Note that it can be due to the governmental social policy in rural areas which aims to hold a higher level of employment. The highest non-radial overuse of machinery was observed in the farm group while the lowest among cooperatives and “other” forms. An explanation for this may be that many farms do not have enough machinery for their needs and therefore hire more machinery services from other companies which, generally, is more expensive.

Technical efficiency under the assumption of non-increasing return to scale was also estimated in order to find out under which return to scale (increasing, constant or decreasing) each enterprise operates. The results of this evaluation are reported in Table 7.

*** **Table 7 here** ***

Firms which operate under decreasing return to scale (DRS) prevail in the regions of Cherkassy and Kiev (45 and 52% respectively). In the region of Poltava, which has a larger average farm size compared to the other two regions, the share is the highest (55%). 1-3 % of all analyzed enterprises in each region were operating under constant return to scale. The rest of farm, corresponding to 42 - 52 percent depending on region, operates under increasing return to scale (IRS). A situation when decreasing return to scale prevails (which was the case for Cherkassy and Kiev) is expected as size of agricultural firms in Ukraine is on average quite big. As can be seen in Table 5, the scale efficiencies are on average quite high for all groups of farms. This could mean that many farms that operate under DRS and IRS are close to “fully” scale efficient. It should be pointed out that we should be very careful with conclusion about most productive scale size as cost-based production possibility set was used in this analysis. Thus, we cannot say that there is some optimal size only with respect to return to scale. This issue will be discussed later in the second stage analysis.

*** **Figure 1 here** ***

It is also interesting to look at the distribution of efficiency. Frequency cumulative distribution histogram of efficiency scores calculated under variable return to scale is presented in Figure 4. We can observe that approximately 80 % of all firms in the analyzed time periods have efficiency scores less than 0.7 and approximately 60 % have less than 0.6. This suggests that the overuse of inputs to produce certain volume of wheat is substantial and leaves room for improvement. To find out what input factors that are particularly overused one can compare shadow shares and actual shares as it is done in Figure 2 below.

*** **Figure 2 here** ***

Figure 2 shows that some disparities between actual and shadow shares exist although these are rather small. Labor was somewhat overused in 2006, less overused in 2007 and used "optimally" in 2008. Variable inputs were on average underused trough all years and disparity between actual and shadow shares in 2008 was even higher than in 2006. Such difference in variable inputs shares is likely explained by a gap between the need and availability of some of them (especially fertilizers).

In order to analyze the impact of various factors on technical and cost efficiency, Tobit (censored) regressions were conducted in which each of the efficiency measures where used as dependent variables. Tobit regressions are the conventional method used for the second stage analysis.[7]

The dependent variables include land used for wheat (*Land*) and share of total land used for wheat (*Swheat*). The squared values for each of these variables where also used as dependent variables (*Land2* and *Swheat2*) as it can be expected that they have are non-linearly related to efficiency (e.g., there may be an "optimal" share of wheat). We further included a variable that indicates the farms specialization in crop production, *SpecC*. This variable is constructed as the share of revenues from crop production of total agricultural revenue. Apart from these variables, dummy variables for legal form, region and year was used as explanatory variables.

*** **Table 8 here** ***

The results suggest that land size has a positive and significant impact on technical efficiency under the assumption of variable returns to scale and cost efficiency. The significant negative values of the coefficient for $Swheat^2$ in all three models suggest that there is some optimal share of land used for wheat cultivation. The level of specialization in crop production is found to be negatively related to efficiency (for all three types). This is likely explained by the fact that firms with no livestock only grow grain and oilseeds which leads to bad crop rotation. An additional

---

[7] Simar and Wilson (2007) argues that since there are an unknown form of serial correlation among the DEA efficiency scores, standard methods causes invalid inference. They suggest two bootstrap procedures that can be used to make valid inference. In our case, the application of these algorithms didn't have any implication for the main findings and we therefore report the results of a conventional Tobit model.

reason may be that livestock producers produce and use organic fertilizers whereas crop farms mainly buy mineral fertilizers that only have a big effect in the short run. The suggestion about higher efficiency of cooperatives is now confirmed as Table 8 demonstrates positive and significant parameter values for all three types of efficiency. Significant parameter estimates were found for the region and year dummy variables.

## 5. Discussion and Conclusions

This study analyzed efficiency in wheat production of agricultural enterprises in three Ukrainian regions (Kiev, Poltava and Cherkasy). The findings suggest that there were substantial variations in efficiency scores among agricultural enterprises during the analyzed time period (2006-08). Thus, the results indicate that there is quite big scope for efficiency improvement in wheat growing among the analyzed enterprises.

Despite prevailed decreasing return to scale in two of the regions (Kiev and Poltava), scale efficiency were on average high. As the average technical efficiency under VRS was low and very close to technical efficiency under CRS and allocative efficiency scores were comparatively high, this suggests that the main source of inefficiency is pure technical inefficiency. However, since inputs in value terms were utilized in our analysis, the VRS technical inefficiency is also likely to be affected by differences in input prices across the firms. Since we do not have sufficient information about prices and physical quantities of used inputs, losses due to price inefficiency cannot be indicated in an accurate way. A more detailed analysis that considers such information (if available) is suggested for future researches in this area.

Concerning to the impact of various factors on the efficiency measures, it was found that location of the farm have an impact on efficiency (with farms in Poltava being most efficient) and that cooperatives are more efficient compared to the other legal forms. Enterprise size (measured by land) was found to have a positive and statistically significant impact on technical (VRS) and cost efficiency. A higher efficiency level was observed in 2008 which can be explained by the beneficial weather conditions that year. Concerning, over-/underuse of inputs and outputs it was found that variable inputs (seed, fertilizers etc) was underused in all years whereas labor input was overused in the first years but used in optimal proportions in year 2008.

## References


Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Some models for estimating technical and scale inefficiences in data envelopment analysis. *Management Science* 30(9): 1078-1092.

Charnes, A. Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research* 2(4): 429-444.

Coelli, T.J., Prasada Rao, D.S., O'Donnell, C.J. and Battese, G.E. (2005). *An introduction to efficiency and productivity analysis.* Kluwer Academic Publishers.

Coelli, T.J. and Prasada Rao, D.S. (2001). Implicit value shares in Malmquist TFP Index Numbers, CEPA Working Papers, No. 4/2001, School of Economics, University of New England, Armidale.

Cooper, W., Seiford, L. and Tone, K. (2007). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Second Edition, Springer.

Cross, R., and Färe, R (2008). Farrell efficiency under value and quantity data. *Journal of productivity analysis* 29: 193-199.

Farrell, M.J. (1957). The measurement of Productive Efficiency, *Journal of the Royal Statistical Society*, Series A, CXX, Part 3, 253-290.

Färe, R. Grosskopf, S. and Lovell, C.A. (1985). *The measeurement of efficiency of production*. Kluwer Academic Publishers.

Färe, R. Grosskopf, S. and Lovell, C.A (1994). *Production Frontiers.* Cambridge University Press.

Lissitsa, A. and Odening, M. (2005). efficiency in total factor productivity in Ukrainan agriculture in transition. *Agricultural Economics* 32: 311-325.

Simar, L. and Wilson, P.W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in non-parametric frontier models. *Management Science* 44: 49-61.

Simar, L. and Wilson, P.W. (2000). Statistical inference in non-parametric frontier models: the state of the art. *Journal of Productivity Analysis* 13(1): 49-78.

Simar, L. and Wilson, P.W. (2007). Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Processes, *Journal of Econometrics* 136: 31-64.

Staat, M. (2002). Bootstrapped efficiency estimates for a model for groups and hiearchies in DEA. *European Journal of Operational Research* 138: 1-8.

Singh, S., Fleming, E. and Coelli, T. (2000). Efficiency and Productivity Analysis of Cooperative Dairy Plants in Haryana and Punjab States of India, *University of New England Working Paper* ISSN 1442 1909.

Tone, K. (2002). A Strange Case of the Cost and Allocative Efficiencies in DEA A Strange Case of the Cost and Allocative Efficiencies in DEA. *The Journal of the Operational Research Society* 53(11): 1225-1231.

Wilson, Paul W. (2008). FEAR: A Package for Frontier Efficiency Analysis with R. *Socio-Economic Planning Sciences* 42(4): 247-254.

## Tables

**Table 1:** Description sample farms according to region and year. Source: Authors.

| Indicators | 2006 | | | 2007 | | | 2008 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Kiev | Poltava | Cherkasy | Kiev | Poltava | Cherkasy | Kiev | Poltava | Cherkasy |
| Number of farms | 203 | 215 | 195 | 237 | 230 | 241 | 257 | 263 | 269 |
| Average farms size in hectares (s.d.*) | 2356 | 2994 | 2359 | 2195 | 2881 | 2411 | 2375 | 2832 | 2435 |
| | (3887) | (2359) | (2523) | (4359) | (3075) | (3317) | (5974) | (3167) | (3195) |
| Average wheat yield per hectare (s.d.*) | 3.08 | 2.76 | 3.2 | 2.91 | 2.93 | 2.99 | 4.02 | 4.47 | 4.70 |
| | (1.16) | (0.74) | (0.93) | (0.98) | (0.80) | (1.10) | (1.33) | (1.14) | (1.15) |
| Share of arable land used for wheat, % (s.d.*) | 26.0 | 18.8 | 18.9 | 26.8 | 23.1 | 21.9 | 26.9 | 24.3 | 22.4 |
| | (12.9) | (10.2) | (9.5) | (12.2) | (11.4) | (11.3) | (12.3) | (12.4) | (9.8) |

*s.d. = standard deviation

**Table 2:** Description of data. Source: Authors.

| Variables | Mean | Standard deviation | Minimum | Maximum |
|---|---|---|---|---|
| **Output variable:** | | | | |
| Production of wheat (m.t.) | 2012 | 4315 | 9.00 | 135049 |
| **Input variables:** | | | | |
| Labor (ths. UAH) | 123 | 282 | 0.88 | 8869 |
| Variable inputs (ths. UAH) | 536 | 1243 | 3.23 | 33697 |
| Machinery (ths. UAH) | 354 | 701 | 2.79 | 16806 |
| Other costs (ths. UAH) | 216 | 603 | 0.13 | 14136 |
| **Specific variables:** | | | | |
| Land used for wheat (ha) | 540 | 871 | 6.00 | 21485 |
| Total farm size (ha) | 2538 | 3735 | 23.00 | 70018.00 |
| Share of total land used for wheat (%) | 23 | 12 | 0.75 | 100.00 |
| Specialization in crop growing [8] (%) | 74 | 24 | 0.21 | 100.00 |

[8] Share of revenue from plant growing in total revenue from agricultural production.

**Table 3:** Legal forms by year. Source: Authors.

| Legal forms | 2006 | | 2007 | | 2008 | |
|---|---|---|---|---|---|---|
| | Number of firms | % | Number of firms | % | Number of firms | % |
| Limited liability company | 419 | 68 | 492 | 69 | 542 | 67 |
| Private enterprise | 135 | 22 | 146 | 21 | 162 | 21 |
| Farm | 4 | 0.6 | 10 | 1 | 19 | 2 |
| Cooperative | 41 | 7 | 40 | 6 | 36 | 5 |
| Other form | 14 | 2 | 20 | 3 | 29 | 4 |

**Table 4:** Description sample farms according to legal form and year. Source: Authors.

| Legal forms | Average per year (standard error within brackets) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 2006 | | | 2007 | | | 2008 | | |
| | Size (ha) | Wheat yield per ha | Share of land used for wheat (%) | Size (ha) | Wheat yield per ha | Share of land used for wheat (%) | Size (ha) | Wheat yield per ha | Share of land used for wheat (%) |
| Limited liability company | 2724 | 3.03 | 19.9 | 2646 | 3.00 | 22.9 | 2709 | 4.48 | 24.2 |
| | (3492) | (0.96) | (11.0) | (4094) | (0.97) | (11.2) | (4908) | (1.25) | (12.2) |
| Private enterprise | 2097 | 2.95 | 23.6 | 2016 | 2.73 | 27.1 | 2073 | 4.20 | 25.6 |
| | (1390) | (1.05) | (10.6) | (2405) | (0.93) | (13.4) | (2738) | (1.18) | (11.0) |
| Farm | 3326 | 3.45 | 20.4 | 2253 | 2.45 | 31.2 | 2246 | 4.14 | 18.3 |
| | (2090) | (1.59) | (7.1) | (1938) | (0.54) | (19.5) | (1516) | (1.46) | (8.5) |
| Cooperative | 2414 | 2.92 | 24.2 | 2177 | 2.91 | 23.6 | 2214 | 4.34 | 25.0 |
| | (1374) | (0.84) | (14.9) | (1416) | (1.05) | (8.3) | (1379) | (1.38) | (8.2) |
| Other form | 3219 | 2.86 | 27.8 | 2900 | 3.37 | 22.7 | 2831 | 4.45 | 26.5 |
| | (1780) | (0.88) | (15.8) | (2459) | (1.09) | (11.0) | (2170) | (0.93) | (9.3) |

**Table 5:** Efficiency scores. Source: Authors.

| Group of farms | | Technical efficiency (CRS) | | | | Technical efficiency (VRS) | | | | Ce† | Ae†† | Se††† |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ***Original*** | ***Bias-corrected*** | ***CI(1)**** | ***CI(2)***** | ***Original*** | ***Bias-corrected*** | ***CI(1)**** | ***CI(2)***** | | | |
| *All* | | 0.57 | 0.53 | 0.51 | 0.56 | 0.59 | 0.54 | 0.52 | 0.58 | 0.51 | 0.92 | 0.96 |
| Region | Kiev | 0.52 | 0.49 | 0.47 | 0.51 | 0.55 | 0.50 | 0.48 | 0.54 | 0.47 | 0.92 | 0.96 |
| | Poltava | 0.61 | 0.57 | 0.55 | 0.60 | 0.63 | 0.58 | 0.55 | 0.62 | 0.56 | 0.93 | 0.96 |
| | Cherkasy | 0.57 | 0.54 | 0.51 | 0.56 | 0.59 | 0.54 | 0.52 | 0.58 | 0.51 | 0.91 | 0.96 |
| Legal form | Limited liability company | 0.56 | 0.53 | 0.51 | 0.55 | 0.59 | 0.5 | 0.51 | 0.57 | 0.51 | 0.92 | 0.96 |
| | Private enterprise | 0.57 | 0.54 | 0.52 | 0.56 | 0.60 | 0.55 | 0.52 | 0.58 | 0.52 | 0.92 | 0.97 |
| | Cooperative | 0.61 | 0.58 | 0.56 | 0.60 | 0.64 | 0.59 | 0.56 | 0.62 | 0.55 | 0.91 | 0.96 |
| | Farm | 0.56 | 0.53 | 0.51 | 0.55 | 0.60 | 0.55 | 0.52 | 0.59 | 0.51 | 0.90 | 0.94 |
| | Other | 0.56 | 0.53 | 0.51 | 0.55 | 0.59 | 0.53 | 0.50 | 0.57 | 0.52 | 0.93 | 0.97 |

* Average lower limit of 90% confidence interval, ** Average upper limit of 90% confidence interval, †Ce = Cost efficiency, †† Ae= allocative efficiency and ††† = Scale efficiency

**Table 6:** Averages and percentages of input slacks. Source: Authors.

| Legal forms | Average, ths. UAH | | | | Percentage, % | | | |
|---|---|---|---|---|---|---|---|---|
| | Labor | Variable inputs | Machinery | Other costs | Labor | Variable inputs | Machinery | Other costs |
| Limited liability company | 5.14 | 4.17 | 3.05 | 3.50 | 1.4 | 0.22 | 0.24 | 0.55 |
| Private enterprise | 2.94 | 1.02 | 3.01 | 4.21 | 1.3 | 0.22 | 0.31 | 0.43 |
| Cooperative | 5.40 | 0.09 | 0.01 | 2.06 | 1.7 | 0.01 | 0.00 | 0.70 |
| Farm | 0.19 | 11.7 | 33.7 | 2.39 | 0.07 | 1.15 | 2.48 | 0.80 |
| Other | 2.17 | 0.98 | 0.00 | 3.52 | 0.34 | 0.07 | 0.00 | 0.55 |

**Table 7:** Results of returns to scale evaluation. Source: Authors.

| Return to scale (RTS) | Kiev | | Poltava | | Cherkasy | |
|---|---|---|---|---|---|---|
| | **Number of firms** | **%** | **Number of firms** | **%** | **Number of firms** | **%** |
| Increasing RTS | 328 | 47 | 298 | 42 | 370 | 52 |
| Constant RTS | 7 | 1 | 19 | 3 | 20 | 3 |
| Decreasing RTS | 362 | 52 | 391 | 55 | 315 | 45 |
| All | 697 | 100 | 708 | 100 | 705 | 100 |

**Table 8:** Efficiency determinants (Tobit regressions). Source: Authors.

| Explanatory variables | | Dependent variable | | |
|---|---|---|---|---|
| | | Technical efficiency CRS (Standard Error) | Technical efficiency VRS (Standard Error) | Cost efficiency (Standard Error) |
| *Constant* | | 0.459*** (0.0173) | 0.488*** (0.0185) | 0.410*** (0.0143) |
| *Land* | | -0.780*10^-5 (0.789*10^-5) | 0.368*10^-4*** (0.116*10^-4) | 0.228*10^-5*** (0.655*10^-5) |
| *Land*$^2$ | | 0.297*10^-9 (0.559*10^-9) | 0.248*10^-8 (0.170*10^-8) | 0.172*10^-10 (0.464*10^-9) |
| *Swheat* | | 0.00453*** (0.00084) | 0.00247 (0.000910) | 0.00397** (0.0006989) |
| *Swheat*$^2$ | | -0.627*10^-4*** (0.108*10^-4) | -0.389*10^-4*** (0.116*10^-4) | -0.570*10^-4*** (0.898*10^-5) |
| *SpecC* | | -0.00052*** (0.000155) | -0.00046*** (0.000166) | -0.00049*** (0.000129) |
| **Dummy variables** | | | | |
| **Legal form** | *Private enterprise* | 0.0813 (0.00897) | 0.0864 (0.00962) | 0.00960* (0.00744) |
| | *Cooperative* | 0.04811*** (0.0159) | 0.0560*** (0.0170) | 0.0430*** (0.01312) |
| | *Farm* | -0.01363 (0.0285) | -0.00398 (0.0305) | -0.0199 (0.0237) |
| | *Other legal form* | -0.00477 (0.021) | -0.0123 (0.0224) | 0.00147 (0.0237) |
| **Region** | *Poltava* | 0.0975*** (0.00896) | 0.0898*** (0.00965) | 0.0956*** (0.00743) |
| | *Cherkasy* | 0.0588*** (0.00893) | 0.0546*** (0.00956) | 0.0514*** (0.00740) |
| **Year** | *2007* | -0.0288*** (0.00901) | -0.0266*** (0.00964) | -0.0228*** (0.00736) |
| | *2008* | 0.107*** (0.00887) | 0.118*** (0.00949) | 0.113*** (0.00736) |

***, **, and * indicates statistical significance on 1, 5 and 10% levels.

**Figures**

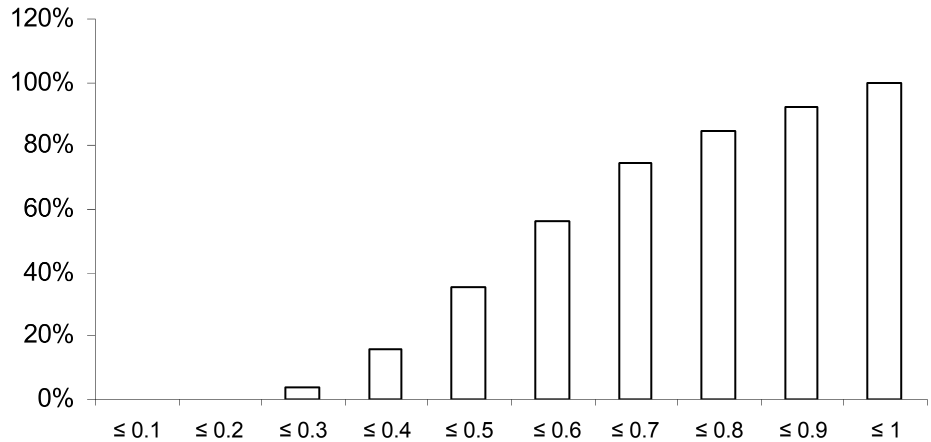


**Figure 1:** Cumulative distribution of technical efficiency scores under VRS. Source: Authors.

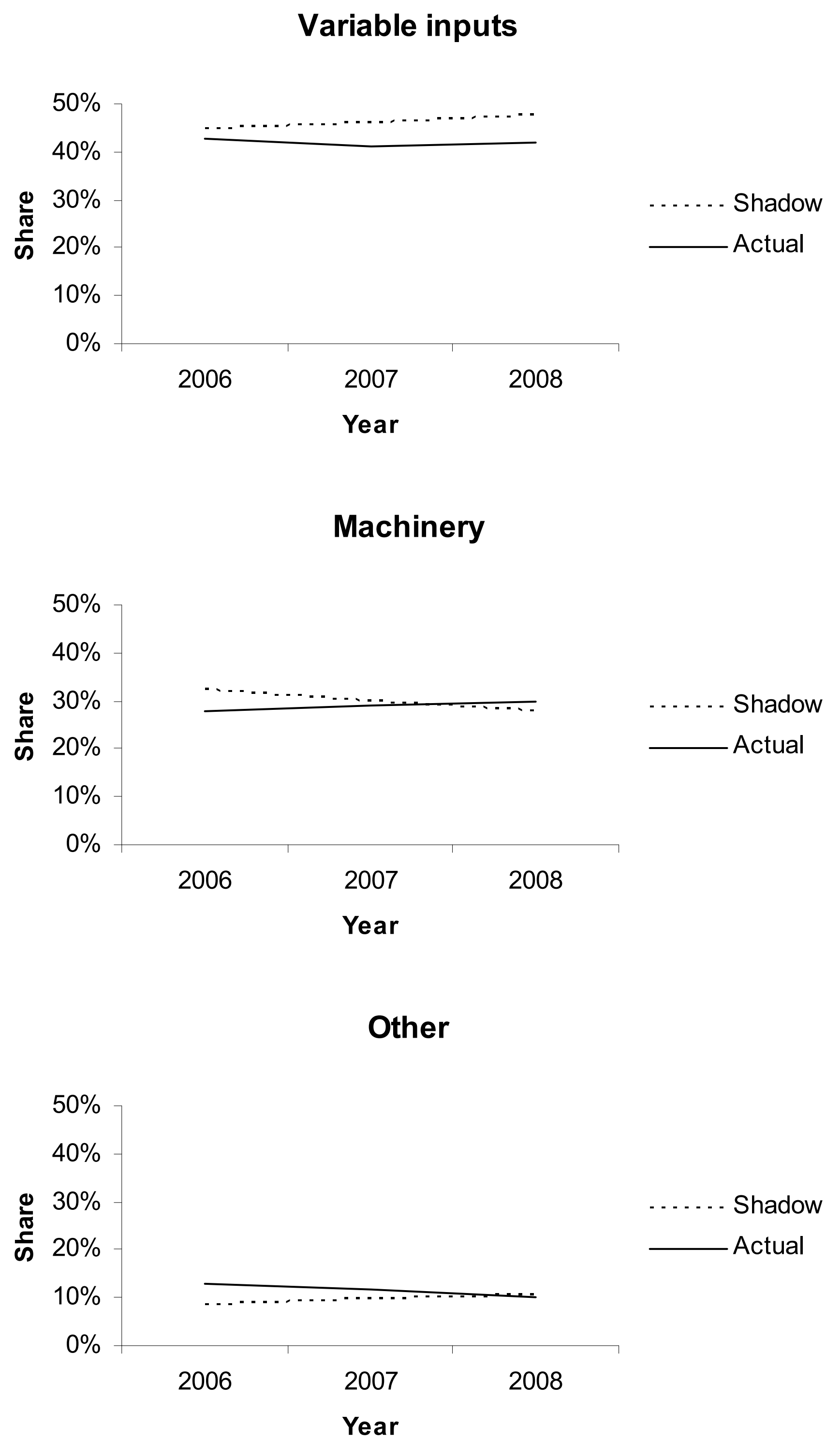


**Figure 2:** Development of actual and shadow value share of inputs. Source: Authors.