# EFFICIENCY CHANGE IN UKRAINE'S SUNFLOWER SEEDS INDUSTRY: 1998-2002

NIVYEVS'KIY OLEG [*], ZELENYUK VALENTIN [**],

**ABSTRACT**

In this paper we measure efficiency of Ukraine's Sunflower Seeds Industry over the period 1998-2002. The efficiency scores are measured on the level of particular sunflower seed producers, then aggregated into and compared on the regional level. All estimations are done using Data Envelopment Analysis approach. We also used methodology for obtaining aggregate efficiencies scores of distinct groups with weights derived from economic optimization, following FÄRE AND ZELENYUK (2003). The statistical inference for the aggregate scores and for the difference between them is done using statistical bootstrap approach, following SIMAR AND ZELENYUK (2003).

**Keywords:** sunflower seeds, efficiency, applied bootstrap, Ukraine.

## 1 INTRODUCTION

In this paper we focus particularly on measuring efficiency of Ukraine's Sunflower Seeds Industry over the period 1998-2002. The efficiency scores are measured on the level of particular sunflower producers (to identify the sources of inefficiency) and then aggregated to industry level. We estimate efficiency using Data Envelopment Analysis (DEA) approach, then inference on group efficiencies using statistical bootstrap proposed by SIMAR AND ZELENYUK (2003).

There are several reasons why we think this study is interesting. First of all, sunflower growing became one of the most profitable industries in Ukraine's agriculture. For example, sunflower seed profitability constituted 77.9 % in 2002, while grain profitability constituted only 19.3% the same year (State Committee of Statistics of Ukraine, 2002). In 2002/03 marketing year (MY), the total output of sunflower seeds of Ukraine constituted 17% of the world total output of sunflower seeds. Moreover, the export of Ukrainian sunflower seeds was 14.8% of the total world sunflower seeds export in 2002\03 MY. This implies strategic importance of industry for domestic sunflower oil producers and Ukraine as a whole. Nevertheless, none of the previous studies looked specifically at measuring efficiency of sunflower seed industry operation. Previous studies covered only the early period of Ukraine's independence (JOHNSON et al., 1994), or looked at the

[*] Institute for Economic Research and Policy Consulting, Ukraine. Email: onivyevskiy@eerc.kiev.ua, nivyevskiy@ier.kiev.ua

[**] UCL, Belgium. Email: vzelenyuk@eerc.kiev.ua

whole agriculture (GALUSHKO et al., 2003), or only at grain industry for some regions of Ukraine (KURLAKOVA AND JENSEN, 1998).*

One of the particular goals of the paper is to shed some light on the effectiveness of 2000 Land reform. Its essence was that state and collective farms were formally disbanded, and the land of collective farms was distributed among the people of those farms. Despite our expectations, we have found no strong evidence for positive influence of 2000 Land reform on efficiency of sunflower seeds industry operation. In general, we found the sunflower seeds production in Ukraine having low efficiency.

## 2 BRIEF INDUSTRY OVERVIEW

### 2.1 Land Reform 2000

One of the latest most prominent interventions of government to Ukrainian agriculture was comprehensive land reform, executed in early 2000.* The pivotal point of the reform was formal disbanding of the Soviet-era state and collective farms whose land must have been distributed among the people of those farms. In reality, a bulk of land was leased back to the old managers of the state and collective farms; a good portion of land went to private plots and enterprises, and to the large commercial holdings (ASLUND, 2002). In particular, four types of enterprises have appeared in Ukrainian agriculture: (i) *Limited liability companies* (Ltd), (ii) *Cooperatives*, (iii) *Private enterprises*, and (iv) *State-owned enterprises*.

As one can see, legally agricultural land ownership became quite dispersed, but what are the consequences to that? There were many talks whether this reform was effective or not on political arena and the media. This paper simply proposes to see "what the data says". In particular, after having estimated individual efficiency scores for fife years for each enterprise, we can test the effect of reform on performance of this particular types of firms by comparing estimated densities of distribution of efficiency scores, as well as comparing group (aggregate) efficiencies using samples representing various types of ownerships.

### 2.2 Demand for Sunflower Seeds

Produced sunflower seeds go mostly to sunflower processing plants or to export. Remarkably, before export duty imposition (23% in 1998), Ukraine was among the three largest exporters of sunflower seeds (in June 2001, the tax was

* An exception is, perhaps, a (non-academic) study-report of TACIS (1999), which in particular made somewhat provocative statement that the rate of capacity utilization has dropped from 80% in 1990 down to 34% in 1998.

* Decree of the President of Ukraine ? 1529/99 "On pressing measures on acceleration of reforming of agrarian sector of economy"

lowed to 17%). Export tax had consequences for internal market and for oil-fat industry. In particular, it resulted in initial sharp reduction in export, increased domestic supply and hence lowered domestic prices. These lower domestic prices impacted the domestic oil-fat industry, reflecting in the increase of domestic oil supply hence in oil export growth (15.88 % of total sunflower oil world export).

Over the last five years demand for sunflower seeds is stable due to constant product requirements from oil processors and export side. There are 19 sunflower processing plants, united in the *Ukroliyaprom Association*, which accounted for more than 80% of vegetable oil in the country in 2002. The rest of the production is done on small-scale processing enterprises with low capacities and outputs. It is worth noting that currently total capacity of Ukrainian oil-fat processors constitute 3.9 mln tons (MT) of sunflower seeds, compared to 4.2 (MT) of 2003 harvest. Basic mass of sunflower seed is processed by oil-fat processors. Moreover, half of domestically produced sunflower oil is exported abroad. And the rest is consumed domestically (Ukraine's internal vegetable oil requirements are about 450,000 to 500,000 tons) suggesting high domestic demand for sunflower seeds for processing needs.

### 2.3 Supply of Sunflower Seeds

World sunflower seeds production has been increasing over the last decade. It was on average 23.5 (MT) in the mid 1990s up to 26.26 MT for 2003/2004 marketing year (MY). In this situation, for example, Ukraine produced less only then Russia in 2000/2001 MY (3.46 MT), and in 2003/2004 MY (4.5 MT). Interestingly, collective agricultural enterprises were the main sunflower seed producers in 1998 (TACIS, 1999), in particular they accounted for 91% in the total output. Individual land plots accounted for 5.1% and private farms accounted for 3.9% in the total volume. After the land reform, the distribution has changed dramatically. During 2001-2002 years, the Ltd companies started dominating in the production of sunflower seeds among enterprises of different ownership structures. They produced more than 50% of total sunflower seeds output in each region. Privately owned enterprises took the second place: they include Private Family Farms and Private Enterprises. Agricultural Cooperatives took the third place, while State-owned enterprises took the last place.

Almost 90% of Ukraine sunflower production is concentrated in the eastern and southern regions; among them the most important are Dnipropetrovs'k (15.13%), Zaporizhia (13.78%), Donets'k (12.17%) and Kharkiv (10.15%). For this reason we will mostly focus our analysis on these regions.

# 3 METHODOLOGY

## 3.1 Technology and Efficiency Characterizations

Assume $n$ firms operate in a region $r$ *(r=1,…,R)* at question. Each firm $k$ *(k=*$1,...,n_r$*)* in region $r$ uses $N$ inputs, denoted with $x^k = (x_1^k,...,x_N^k)' \in \Re_+^N$, to produce $M$ outputs, denoted with $y^k = (y_1^k,...,y_M^k)' \in \Re_+^M$. We assume that within a region, all $n$ firms have access to the same technology $T^r$, defined in general terms as

$$T^r \equiv \{(x\ ,y\ ) : x \quad can \quad produce \quad y\ \}, \qquad r = 1,...,R \tag{1}$$

that satisfies standard regularity axioms of production theory (e.g. see FÄRE, GROSSKOPF AND LOVELL (1994). Technology is allowed to differ between regions. Under these axioms we can use the *output oriented* SHEPHARD'S (1970) distance function $D_o^r : \Re_+^N \times \Re_+^M \to \Re_+^1 \cup \{\infty\}$, defined as

$$D_o^r(x\ ,y\ ) \equiv \inf\{q : \ (x\ ,y\ /q) \in T^r\}, \tag{2}$$

to completely characterize technology set $T^r$ of region $r$. This distance function can be used to define *Farrell*-type output oriented technical efficiency measure for firm $k$

$$TE^r(x^k, y^k) \equiv 1 / D_o^r(x^k, y^k) \tag{3}$$

Whenever we state that $D_o^r(x^k, y^k) = 1$ or $TE^r(x^k, y^k) = 1$, we assert that firm $k$ is *technically* efficient relative to frontier of region $r$, otherwise, when $TE^r(x^k, y^k) > 1$, it is *technically* inefficient. For convenience, one can represent efficiency score of a firm $k$ in percentages, i.e. $\left(1/TE^r(x^k, y^k)\right)*100\%$ and its inefficiency score would then be $\left(1-1/TE^r(x^k, y^k)\right)*100\%$.

## 3.2 Aggregation

While aggregating individual efficiencies into the sub-group or group levels within a region, naturally, one may want to account for *contribution* of particular firm in total group (region) score. Since we will consider aggregation within region, we will drop superscript $r$ for simplicity of notation in this section. Suppose we want to estimate aggregated efficiency of sub-group $l$ ($l = 1,...,L$) within some group, represented by sub-sample $k = 1,...,n_l$. FÄRE AND ZELENYUK (2003) and SIMAR AND ZELENYUK (2003) have proposed aggregating individual efficiency scores into group efficiencies with weights coming from economic optimisation criterion. In particular, their aggregate efficiency for *sub-group l* is obtained as

$$\overline{TE}^l \equiv \sum_{k=1}^{n_l} TE\ (x^{l,k}, y^{l,k}) \cdot S^{l,k}, \ S^{l,k} \equiv \frac{py^{l,k}}{p\bar{Y}^l}, \ \bar{Y}^l = \sum_{k=1}^{n_l} y^{l,k}, \qquad l = 1,...,L \tag{4}$$

By the same manner, the aggregate efficiency score for entire group (i.e., aggregated over all the sub-groups of a region) is given by

$$\overline{TE} = \sum_{l=1}^{L} \overline{TE}^{l} \cdot S^{l}, \qquad S^{l} = p\overline{Y}^{l} / p\sum_{l=1}^{L}\overline{Y}^{l} \tag{5}$$

Particularly for our study, this formula tells us that technical efficiency for region's production of sunflower seeds is obtained by averaging the group efficiencies over all enterprises (regions) with weights being the shares of each firm in total regional revenue. It is worthwhile to note that for a single-output case (as ours), the weights become the corresponding output-shares, and the aggregate measure becomes what FARRELL (1957) called as *Individual Structural Efficiency*.

### 3.3 DEA Estimation

The technology set $T^r$ for each region *r (r=1,…,R)* is unobservable but can be (under certain assumptions[*]) consistently estimated using *Activity Analysis Model*, with the following set

$$\hat{T}^{r} \equiv \Big\{ (x, y): \; y \le \sum_{k=1}^{n_r} z_k y_m^k; x \ge \sum_{k=1}^{n_r} z_k x_i^k; \sum_{k=1}^{n} z_k = 1; z_k \ge 0; \quad k=1,...,n_r; m=1,...,M; i=1,...,N; r=1,...,R \Big\} \tag{6}$$

The boundary of this set defines what is called the *observed 'best-practice frontier'*. Such approximation of the true technology can be done for each region and then the individual efficiency can be estimated relative to the observed best practice frontier of the corresponding region. In particular, solving the following linear programming problem

$$\hat{TE}^{r}(x^j, y^j) \equiv \max_{\Theta, Z_1,...,Z_K} \Big\{ \Theta: \; \Theta y^j \le \sum_{k=1}^{n_r} z_k y_m^k; x^j \ge \sum_{k=1}^{n_r} z_k x_i^k; \sum_{k=1}^{n_r} z_k = 1; z_k \ge 0; \quad k=1,...,n_r; m=1,...,M; i=1,...,N \Big\} \tag{7}$$

for each observation (firm) $j = 1,...,n$ in the sample gives estimate of technical efficiency for the particular firm *j*. The estimates of aggregate efficiency scores are obtained by replacing (in formulas *(3)* and *(4)*) the unknown individual technical efficiency scores with their DEA estimates.

### 3.4 Bootstrap for DEA

Since in DEA we obtain the best practice frontier from the observed data, it must be clear that $\hat{T}^r \subseteq T^r$. So, $\hat{TE}^r(x, y)$ is a downward biased estimator of $TE^r(x, y)$, i.e., $1 \le \hat{TE}^r(x,y) \le TE^r(x,y), \forall \; (x,y) \in \hat{T}^r(x,y)$. KNEIP ET AL (1998) have proved consistency of the DEA estimator, while KNEIP ET AL (2003) have derived its limiting distribution and proved consistency of their algorithm of sub-sampling bootstrap for DEA. SIMAR AND ZELENYUK (2003) have extended the bootstrap for

[*] See KNEIP ET AL (2003), and SIMAR AND ZELENYUK (2003) for details

the case of inference on aggregate efficiencies in group-wise heterogeneous context. Let us briefly outline the essence of their approach we will use here.

Clearly, we cannot observe aggregate efficiencies outlined in section 3.2, but we can estimate them replacing true individual efficiency scores in formulas *(4)* and *(5)* with their consistent DEA estimates $\widehat{TE}^r(x^k, y^k)$. So we can get estimate of our true group efficiency score $\overline{TE}^{l,r}$, i.e. $\overline{\widehat{TE}}^{l,r}, l=1,...,L$, for each region *r*. We are interested in the sampling distribution of $\overline{\widehat{TE}}^{l,r} - \overline{TE}^{l,r} \mid \wp$, where $\wp$ stands for data generating process for the distribution of our observed data, $\Xi_n = \{(x^k, y^k) \ : \ k = 1, ..., n\}$. The basic idea of bootstrap is to approximate this distribution by the distribution of $\left\{ \overline{\widehat{TE}_b}^{*l,r} - \overline{\widehat{TE}}^{l,r} \mid \hat{\wp} : b = 1,...,B \right\}$, where $\hat{\wp}$ is a consistent estimator of $\wp$ and $\overline{\widehat{TE}_b}^{*l,r}$ is a bootstrap analogue of $\overline{\widehat{TE}}^{l,r}$ for bootstrap replication *b = 1, …, B*. These bootstrap analogues are obtained by applying the same formula for pseudo-samples $\Xi^*_m = \{(x^{*k}, y^{*k}) : k = 1, \ ..., \ m, \quad m \le n \}$, which are drawn randomly (with replacement) from the true sample $\Xi_n$ (for details and algorithm see SIMAR AND ZELENYUK (2003)). If the bootstrap is *consistent* then we get the following result

$$\overline{\widehat{TE}}^{l,r} - \overline{TE}^{l,r} \mid \wp \overset{asy}{\sim} \overline{\widehat{TE}}^{*l,r} - \overline{\widehat{TE}}^{l,r} \mid \hat{\wp} \qquad (8)$$

i.e., the relationship between the true and estimated (aggregate) efficiency score is approximated by the relationship between the estimated (aggregate) efficiency score and its bootstrap analog. Using this relationship we can consistently estimate the bias and confidence intervals for $\overline{\widehat{TE}}^{l,r} \mid \hat{\wp}$. In particular, given the true bias is

$$Bias\left(\overline{\widehat{TE}}^{l,r} \mid \wp\right) = E\left(\overline{\widehat{TE}}^{l,r}\right) - \overline{TE}^{l,r} \qquad (9)$$

the estimate of bias for aggregate efficiency of a group *l* (*l*=1,…,L) can be approximated with its bootstrap analog

$$\widehat{Bias}\left(\overline{\widehat{TE}}^{l,r} \mid \hat{\wp}\right) = \frac{1}{B}\sum_{b=1}^{B} \overline{\widehat{TE}_b}^{*l,r} - \overline{\widehat{TE}}^{l,r} \qquad (10)$$

where *B* stands for the quantity of samples $\hat{\wp}^*_b$ generated from original sample $\hat{\wp}$. Also true confidence intervals, given by

$$\Pr\left(-b_{a/2} \le \overline{\widehat{TE}}^{l,r} - \overline{TE}^{l,r} \le -a_{a/2} \mid \wp\right) = 1 - a \qquad (11)$$

can be approximated with its bootstrap analog

$$\Pr\left(-\hat{b}_{a/2} \le \overline{\widehat{TE}}^{*\,l,r} - \overline{TE}^{\,l,r} \le -\hat{a}_{a/2} \mid \wp \right) = 1-a \qquad (12)$$

where a is the significance level chosen by researcher and values $\hat{b}_{a/2}$ and $\hat{a}_{a/2}$ are found by sorting $\overline{TE}_b^{*\,l,r} - \overline{TE}^{\,l,r}$, $b = 1,...,B$, and then deleting $\left(\frac{a}{2}\times 100\%\right)$ of the elements at each end, and then setting $\hat{b}_{a/2}$ and $\hat{a}_{a/2}$ to be the endpoints (so that $\hat{b}_{a/2} = \hat{a}_{a/2}$) of this truncated list. Therefore, the resulting bootstrap confidence interval around the unknown aggregate efficiency, $\overline{TE}^{\,l,r}$, is:

$$\overline{TE}^{\,l,r} + \hat{a}_{a/2} \le \overline{TE}^{\,l,r} \le \overline{TE}^{\,l,r} + \hat{b}_{a/2} \qquad (13)$$

### 3.5 Kernel Density Estimation

In this paper we also estimate unknown distributions of random variables — efficiency scores. For this, we use ROSENBLATT (1956) kernel-based estimate of unknown univariate density function $f(u)$, of a random variable $u$, from the sample or realizations of $u$, $\{u_j : j = 1,...,n\}$, defined as

$$\hat{f}(u) = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{u_j - u}{h}\right),$$

where $K$ is a kernel function (e.g. Gaussian density in our case) and $h$ is the bandwidth[*]. The obtained density estimates can be plotted for visual presentation and comparison of distributions.

We will also perform a *formal* comparison of distributions of any two groups, say some group A vs. group Z, using the test statistic proposed by LI (1996) and adapted to DEA context by SIMAR AND ZELENYUK (2004). The hypothesis in this case would be

$$H_O : f_A(u_A) = f_Z(u_Z);$$
$$H_A : f_A(u_A) \ne f_Z(u_Z);$$

where $f_A$ and $f_Z$ are the true (but unknown) density functions of true (unknown, but DEA-estimated) efficiency scores of groups $A$ and $Z$, respectively.

## 4 DATA AND EMPIRICAL RESULTS

This study employs the data on inputs and output of sunflower seed production for regions in Ukraine over the 1998-2002 period. The data at the enterprise level is provided by State Statistics Committee of Ukraine (Table 1).

[*] For estimating and plotting densities we used advanced method for bandwidth estimation proposed by SHEATHER AND JONES (1991). For computation reasons in bootstrap based test of comparing densities we used SILVERMAN (1986) adaptive rule of thumb.

**Table 1. Data Description before (after) Jackstrap Procedure for Selected Regions**

| Regions | | 1999 GVI, UAH thd. | 1999 Arable Land, ha | 1999 Output, 100 kg | 2002 GVI, UAH thd. | 2002 Labor, thd. man/hour | 2002 Output, 100 kg |
|---|---|---|---|---|---|---|---|
| Dnipropet-rovs'k | Obs | 438 *(419)* | 438 *(419)* | 438 *(419)* | 421 *(406)* | 421 *(406)* | 421 *(406)* |
| | Mean | 193.8 *(140.4)* | 599.9 *(601.2)* | 6785.2 *(6542.7)* | 313.9 *(134.4)* | 17.9 *(18.3)* | 7242.4 *(7182.1)* |
| | Std. Dev. | 163.7 *(117.8)* | 416.3 *(410.7)* | 6346.7 *(5955.3)* | 356.02 *(151.35)* | 20.7 *(20.8)* | 7209 *(7194.9)* |
| Donets'k | Obs | 476 *(454)* | 476 *(454)* | 476 *(454)* | 412 *(393)* | 412 *(393)* | 412 *(393)* |
| | Mean | 152.2 *(109.8)* | 459.9 *(463.7)* | 5364.9 *(5185.2)* | 287.24 *(124.39)* | 23.2 *(23.8)* | 6188.1 *(6225.6)* |
| | Std. Dev. | 149.4 *(106.9)* | 370.9 *(368.9)* | 5844.1 *(5495.8)* | 592.9 *(254.5)* | 38.3 *(39.0)* | 11085 *(11257)* |

*Source*: State Statistic Committee of Ukraine

*Notes*: UAH = abbreviation for currency of Ukraine (USD 1=UAH 5.31); numbers after Jackstrap procedure are given in parenthesis

Data is available particularly for the sunflower seeds production and includes *Gross Value of Inputs (GVI)* (for 1998-2002, UAH, in thd), *Land* (for 1998-2000, in ha), *Labour* (for 2001-2002, man-hour, in thd), *Output* (for 1998-2002, in 100 kg). As a result, *to employ as much information as available,* with the data at hand we have to consider two different models for two different periods*. First model operates with 1998-2000 period, using GVI, Land as inputs, and one Output. The second model deals with 2001-2002 period, using GVI, value of Labour, and the same one Output. This happened because of 2000 Land reform, since then it became possible to evaluate land. GVI in the former period includes cost of Labour while the latter model includes value of Land and cost of Labour. Despite including Labour as separate input into the second model, cost of Labour in GVI will allow us to proxy the *quality* of that Labour. Thus, strictly speaking, we should be very careful comparing directly efficiencies estimated from two models, say, for 2002 and 1998 years.

Before referring to results it is important to recall that individual efficiency scores are calculated for each region *separately*, i.e. each observation is measured rela-

* This is one way to address the problem and of course other ways are possible but at this stage are not considered in the paper.

tive to the estimated best-practice frontier of its particular *region*. This is done in order to account for possible heterogeneity across regions (e.g., it is often argued that agroclimatic conditions for growing sunflower are different for regions so measuring all firms with respect to the same frontier for all regions might be in some sense 'unfair').

Since DEA might be very sensitive to the presence of outliers (due to, say, mistyping, misreporting etc), first step we took before estimating efficiency scores was to test for the presence of outliers. We applied the "Jackstrap" procedure proposed by STOŠIC, B. AND SOUSA (2003) and cleared data set from thus detected outliers (see Table 1 for number of deleted items in parenthesis).

### 4.1 Evolution of Efficiency in the Regions

After clearing outliers from the sample using "Jackstrap" procedure of STOŠIC, B. AND SOUSA (2003) we applied the advanced DEA techniques described in the methodology section. Specifically, we first applied smooth homogeneues bootstrap approach of SIMAR AND WILSON (1998) to correct for the bias in original DEA estimates for each firm. Then we estimated the densities of the true efficiency scores using kernel density estimator on the bootstrap-bias-corrected estimates of individual efficiency estimates. Figures 1-2 present the kernel estimated densities for fife years for each region. From simple visual inspection it is hard to infer whether the true densities are different between years for the two major regions. But after applying LI (1996) test adapted for DEA-context (SIMAR AND ZELENYUK, 2004) we can infer that, the true densities for regions under interest differ significantly (Table 2).

For example, for Dnipropetrovs'k region for 1998-1999 years we reject the hypothesis of equality of distributions at 5% significant level. The same result

**Table 2. Li-test results for equality of efficiency distributions across time**

| Null Hypothesis | Dnipropetrovs'k region | | Donets'k region | |
|---|---|---|---|---|
| | ? of obs. for the first (second) group. | p-value | ? of obs. for the first (second) group. | p-value |
| $f(eff_{1998})=f(eff_{1999})$ | 410 (419) | 0.045 | 446 (454) | 0.09 |
| $f(eff_{1999})=f(eff_{2000})$ | 419 (382) | 0.01 | 454 (450) | 0 |
| $f(eff_{2001})=f(eff_{2002})$ | 391 (406) | 0.305 | 427 (393) | 0 |
| $f(eff_{1998})=f(eff_{2002})$ | 410 (406) | 0 | 446 (406) | 0 |

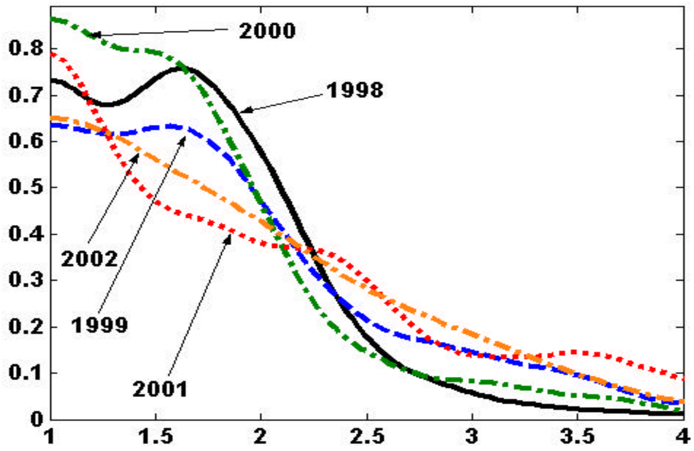


**Figure 1.** Kernel Estimated Densities for Bias Corrected Efficiency Scores. Dnipropetrovs'k 1998-2002

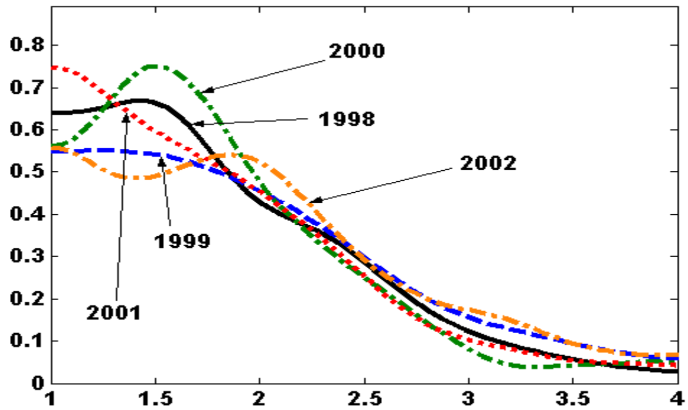


**Figure 2.** Kernel Estimated Densities for Bias Corrected Efficiency Scores. Donets'k 1998-2002

holds for 1998 vs. 2002*, 1999 vs. 2000 years (at 1% significance level), indicating significant changes in firms' performance between the years.

Nonetheless in Figure 2 one may see a mode greater than unity. Intuitively, this means that the tendency in the group of firms was not to be fully efficient, but rather maintain on certain level of inefficiency representing by this non-unity mode called "pathological inefficiency" (SIMAR AND ZELENYUK, 2004). This situation is in contrary to the perfect competition case, where the tendency must be to strive for full efficiency (unity score). In 2000 we observe 'improvement' in distribution versus 1999 distribution; moreover, positive mode reduced that year.

* Since models for 1998 and 2002 are different, strictly speaking, we cannot claim that distributions are really statistically different. Nevertheless, the magnitudes of aggregate scores in two periods (will be discussed below) and behaviour of densities for two periods intuitively leads us to such a conclusion.

In 2001 more mass of density tends to be in a more efficient part but with the grater chance or risk to be inefficient (thick tail). Above all, there was no significant change in 2002 distribution versus the one in 2001. Interestingly, if we compare visually 1998-2002 distributions (although they are significantly different), we notice that 1998 (before the land reform) distribution suggests less inefficiency relative to 2002 distribution (it has more thin tail), witnessing for deterioration of sunflower seeds production, despite the fact that reform happened in between those years. The situation is a bit different for Donets'k region. Here the tails for all densities have visually almost the same thickness. So let us scrutinize the shape of densities closely. As in Dnipropetrovs'k region, there is positive mode in 1998 distribution then follows some tendency to be efficient in 1999. In 2000, density behaves very similar to 1998 density with even greater tendency to have the pathological inefficiency mode. In 2001, density behaves 'nicely' showing firms to be inclined toward the unity efficiency (no inefficiency). 2002 density contrasts with 2001 density sharply by appearance of the non-unity mode[*].

Thus, pattern of distributions of individual efficiency scores witnesses against any substantial, one-way improvements in relative efficiency of sunflower seeds production. On the opposite: it witnesses even in some cases for worsening of relative efficiency of sunflower production in Ukraine from 1998 to 2002. The distributions of efficiency scores fluctuated from one year to another in both directions (improvement and deterioration). Noteworthy, we emphasise that we study the *relative* efficiency between the years: i.e., it could be that some firms have improved over the years so much that the *relative* efficiency of many other firms became very low, relative to what has been originally. Later we will investigate efficiencies of particular groups (types of ownership) of firm.

Now, let us refer to the aggregate efficiency scores. Table 3[**] presents aggregate estimates. In general, the most "influential" regions in production of sunflower seeds are around 55% efficient (Donets'k, Dnipropetrovs'k, and Zaporizhzhia regions) suggesting about relatively low *industry* efficiency, slightly above 50% (asterisk means that respective 95% confidence intervals of estimates overlap). For example, for Dnipropetrovs'k region we observe that aggregate efficiency (bias-corrected) rises from 56,3% in 1998 to 59,36% in 2000, but this change is not significant since 1998 and 2000 95% confidence intervals overlap considerably. Moreover bias corrected estimates for 2000 falls within 1998 confidence interval ('+' means that bias-corrected aggregate efficiency falls within 95% confi-

[*] For other regions also there is no clear-cut tendency in the intertemporal movement of efficiency densities. Moreover there even appeared a pathological inefficiency mode in 2001-2002 for regions like, Zaporizhzhia or Kharkiv.

[**] For the sake of brevity, we present only the first and the last years of the periods considered. But the rest of the estimate can be obtained upon request.

dence interval of another year). The same situation is observed for other regions, except some regions (say, Kharkiv in the Table 3). Nevertheless, those regions aggregate efficiencies do not differ from the main group tremendously (only by about 5-6%). So we cannot really say that those regions developed new technology, had superior position in something, etc. The relatively low aggre

**Table 3. Aggregate Efficiency Scores**

| Regions | 1998 | | | | | 2000 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | AggEff | Bias Corr. AggEff | 95% Conf. Interval Low | 95% Conf. Interval Up | Share | AggEff | Bias Corr. AggEff | 95% Conf. Interval Low | 95% Conf. Interval Up | Share |
| Dnipropetrovs'k*+ | 62.10 | 56.30 | 52.80 | 60.50 | 14.99 | 65.68 | 59.36 | 55.73 | 64.05 | 12.83 |
| Donets'k*+ | 61.40 | 55.20 | 51.90 | 59.60 | 12.04 | 62.58 | 55.22 | 52.02 | 59.66 | 11.47 |
| Zaporizhzhia*+ | 62.10 | 56.50 | 52.90 | 60.90 | 13.62 | 64.57 | 58.81 | 55.45 | 63.02 | 12.01 |
| Kharkiv* | 57.70 | 51.20 | 47.60 | 56.20 | 9.97 | 63.65 | 57.42 | 54.27 | 61.31 | 12.20 |
| **Average for All Regions*+** | 57.03 | 50.49 | 46.68 | 56.15 | | 61.80 | 55.43 | 52.05 | 59.88 | |
| | **2001** | | | | | **2002** | | | | |
| Dnipropetrovs'k*+ | 58.22 | 52.05 | 48.27 | 57.70 | 14.14 | 60.20 | 53.60 | 50.00 | 58.40 | 15.13 |
| Donets'k*+ | 64.37 | 57.93 | 53.74 | 63.26 | 12.83 | 61.30 | 54.50 | 49.70 | 60.60 | 12.70 |
| Zaporizhzhia*+ | 59.28 | 52.65 | 49.39 | 57.03 | 12.70 | 56.90 | 50.30 | 46.90 | 54.90 | 13.78 |
| Kharkiv*+ | 52.73 | 44.64 | 41.13 | 49.03 | 11.74 | 55.90 | 48.40 | 45.00 | 52.90 | 10.15 |
| **Average for All Regions*+** | 55.76 | 48.75 | 44.98 | 53.86 | | 59.878 | 52.374 | 48.11 | 59.77 | |

*Notes: Low* and *Up* stand for low and upper bounds of estimated 95% confidence intervals for the estimates of aggregate efficiency of each region. *AggEff*=aggregate efficiency; *Bias Corr. Agg. Egg.* = Bias Corrected Aggregate Efficiency. For convenience, estimates are presented in $\left(1/\overline{TE}\right)$*100% form.

gate efficiency scores could be explained by the fact that some sunflower producers violated recommendations of proper land rotation frequency for seed. According to the agriculture science sunflower should be sown only once every fife years on the same area. On the contrary, many producers have been doing it every two years (FAO, 2002). As a consequence land was significantly eroded resulting in the yield decline.

The same conclusion may be drawn for 2001-2002 period. Therefore region's aggregate efficiency scores support our conclusions from analyzing distributions of efficiency scores. Thus, received efficiency scores and pattern of their distribution in 1998-2002 years give no evidence for improvement of production efficiency in sunflower seed industry operation. Since there are no analogous estimates for any other crop to compare with (for the time period considered in research) we can at least have a rough perception of our results by comparing them with results of similar research that covered other period. For example, JOHNSON ET AL., (1994) analysed agriculture during 1986-1991 and obtained the estimate of mean efficiency score for grain, using Stochastic Frontier Analysis, was 84%. KURLAKOVA AND JENSEN (2002) estimated mean efficiency score for grain was 71% for the same time period, although they used different model. GALUSHKO ET AL (2003) report that most of enterprises were 50%-75% efficient in 1998-2002 years. Although they used more recent data, DEA approach, theses estimates are for the whole agriculture. Our estimates are lower then scores described above. The reason for that is hidden in data. The model used in our paper concentrates mostly on gross value of inputs (because this is all we have particularly for sunflower seed) whereas GALUSHKO ET AL (2003) used disaggregated data. As FÄRE ET AL (2002) showed aggregation over inputs and/or outputs causes efficiency estimator to be downward biased, i.e. more inefficiency is observed. Moreover they used different model (four inputs one output) thus explaining our low efficiency scores.

## 4.2 Does the Type of Ownership Matter?

Let us concentrate on operation of agriculture enterprises in Dnipropetrovs'k and Donets'k regions in 2001-2002 period—period after the dramatic change in the agricultural ownership structure have emerged with the Land reform. These are the main regions in sunflower seeds industry. Since 2000, the Land reform has led to the radical diversification of ownership, and we are interested in investigating the relative efficiency of sunflower-seed producing enterprises by different ownership structures, in particular by: (i) *Limited Liability Companies* (Ltd.), (ii) *Private agricultural companies* (Private), (iii) *Agricultural cooperatives* (Coop), and (iv) *State-owned agricultural companies*. As was mentioned in the industry overview, Ltd companies produced the largest amount of sunflower seeds in periods under consideration. Private enterprises took the second place, and the state-owned farms produced the least amount of the output.

Table 4 produces estimates for aggregate (group) technical efficiencies for enterprises of the four types *within* each region. Estimates do not produce convincing evidence of better operation in favour of any group of firms. They show that although Ltd companies are the largest producers of sunflower seeds, they appeared to be, on aggregate, as efficient as private companies, cooperatives, and state-owned ones. For example (using confidence intervals criteria), for Dnipropetrovs’k region (2001) state owned enterprises were significantly less efficient than other groups, while Ltd enterprises were as efficient as private enterprises and cooperatives. But the situation has changed in 2002. State-

**Table 4. Aggregate Efficiency Scores for each Group of Enterprises**

| Region | Weighted Aggregate Efficiency scores | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ltd. | | | Private. | | | Coop | | | State | | |
| | Low | Bias Corr. Agg. Eff. | Up | Low | Bias Corr. Agg. Eff. | Up | Low | Bias. Corr. Agg.Eff. | Up | Low | Bias. Corr. Agg.Eff. | Up |
| **Dnipropetrovs'k, 2001** | 51.28 | 52.91 | 54.35 | 52.08 | 54.05 | 55.25 | 52.08 | 53.48 | 54.64 | 40.32 | 42.74 | 44.64 |
| **Donets'k, 2001** | 56.50 | 57.80 | 59.52 | 50.76 | 52.36 | 54.35 | 50.76 | 51.55 | 52.91 | 51.02 | 52.36 | 53.48 |
| **Dnipropetrovs'k, 2002** | 50.51 | 51.81 | 53.19 | 54.35 | 55.87 | 57.14 | 54.95 | 55.87 | 57.14 | 53.76 | 56.18 | 58.14 |
| **Donets'k, 2002** | 51.81 | 53.76 | 55.87 | 48.08 | 49.75 | 51.28 | 51.28 | 53.19 | 54.95 | 51.81 | 53.19 | 54.35 |

*Notes: Low* and *Up* stand for low and upper bounds of estimated 95% confidence intervals for the estimates of aggregate (weighted mean) efficiency of each group. *Bias Corr. Agg. Egg.* = Bias Corrected Aggregate Efficiency. For convenience, estimates are presented in $\left(1/\overline{TE}\right)$*100% form

owned enterprises’ aggregate efficiency score in 2002 was already significantly better than Ltd companies.

For Donets’k region (2001) Ltd enterprises performed significantly better than other groups of enterprises, but they did not do so in 2002. For the rest of the regions, the state-owned enterprises were as efficient as others for the same time period. After all, the pair-wise bootstrap intervals of the RD* (see SIMAR AND ZELENYUK, 2003) statistic reveals that aggregate efficiency scores of different groups are not different from each other. This leads to the conclusion that on the group level, enterprises of different ownership structure performed similarly.

* To save place we do not present results here but they can be provided upon request.

Let us go deeper and consider the distribution of efficiency scores of each group of firms for every region. Figures 3-4 demonstrate kernel estimated densities for

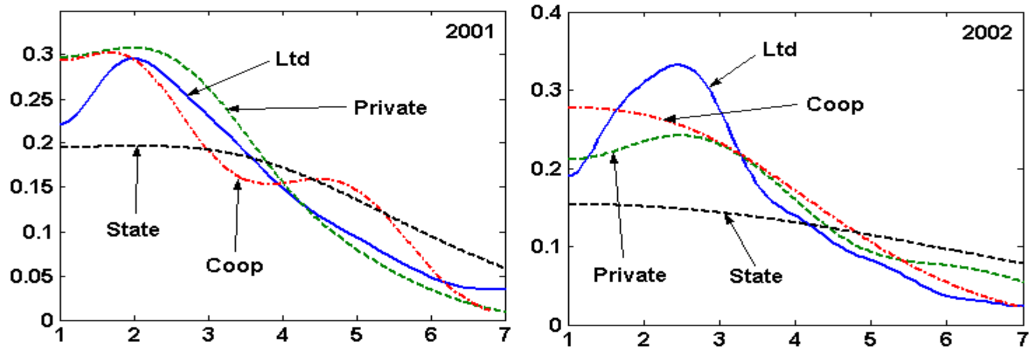


**Figure 3.** Kernel Estimated Densities from Technical Efficiency Scores for Enterprises of Different Ownership Structure, Dnipropetrovs'k. Gaussian kernel is used with the bandwidth selected via SILVERMAN (1986) rule.

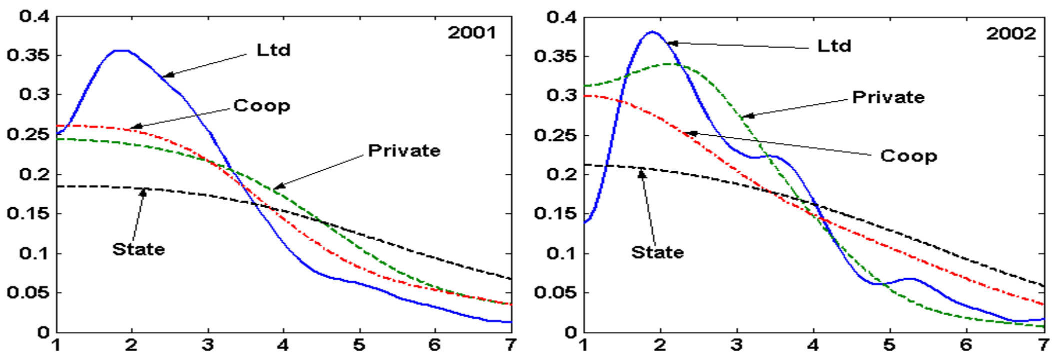


**Figure 4.** Kernel Estimated Densities from Technical Efficiency Scores for Enterprises of Different Ownership Structure, Donets'k. Gaussian kernel is used with the bandwidth selected via SILVERMAN (1986) rule.

firm's efficiency scores. Remarkably, LI (1996) test adapted for DEA-context

(SIMAR AND ZELENYUK, 2004) showed (see Table 5) that pattern of kernel estimated densities between different groups of enterprises does not differ significantly. For example, for Dnipropetrovs'k and Donets'k (2002) regions we cannot reject the hypothesis of equality of densities for different groups of enterprises even at 10% significance level. Nevertheless the patterns of distribution can give some evidence on superiority in efficiency of operation of any particular group vs. another group. Densities for Ltd enterprises for Dnipropetrovs'k and Donets'k regions in 2001-2002 shows 'pathological inefficiency' of operation. We can observe it in a form of a non-unity mode in all distributions for Ltd group on all panels presented. This means that Ltd enterprises had tendency toward the some particular level of inefficiency (on our panels it is around two or 50% efficiency). Hence we have got an empirical support for the hypothesis asserted by ASLUND (2002) that in reality, after 2000 land reform, a bulk of land was leased back to the old managers of the state and collective farms and to the large commercial holdings, thus not bringing much

**Table 5. Li-test results for equality of efficiency distributions for different groups of enterprises. 2002**

| Null Hypothesis | Dnipropetrovs'k. 2002 | | Donets'k. 2002 | |
|---|---|---|---|---|
| | ? of obs. for the first (second) group | p-value | ? of obs. for the first (second) group. | p-value |
| $f_{Ltd}(eff_{Ltd})=f_{Priv}(eff_{Priv})$ | 342 (37) | 0.484 | 313 (62) | 0.62 |
| $f_{Priv}(eff_{Priv})=f_{Coop}(eff_{Coop})$ | 37 (15) | 0.255 | 62 (22) | 0.93 |
| $f_{Coop}(eff_{Coop})=f_{State}(eff_{State})$ | 15 (12) | 0.815 | 22 (15) | 0.45 |
| $f_{Ltd}(eff_{Ltd})=f_{Coop}(eff_{Coop})$ | 342 (15) | 0.035 | 313 (22) | 0.30 |
| $f_{Ltd}(eff_{Ltd})=f_{State}(eff_{State})$ | 342 (12) | 0.79 | 313 (15) | 0.67 |

improvement in performance.

In terms of 'pathological inefficiency' Cooperatives performed the best. All four panels shows cooperatives to be inclined toward the unity efficiency without any non-unity modes, except for Dnipropetrovs'k 2001 (they had some non-unity mode). But in 2002 cooperatives showed improvement in the operation. Private enterprises performed somewhat in between Ltd and Cooperatives in 2001-2002. They also had a 'small' positive mode but not that explicit as Ltd had. State enterprises had no non-unity modes, their distributions resemble a half-normal distribution with high variance, thus suggesting high risk (thick tail) to be very inefficient. Therefore the pattern of distributions showed Cooperatives and Private enterprises having tendency to overperform other groups.

Thus, analysis of group (aggregated) efficiency scores coupled with analysis of distributions of individual efficiency scores between each type of ownership structures fails to present convincing evidence that a particular type of ownership structure of firm producing sunflower seeds has tendency to be more efficient than another type on the aggregate level. But pattern of groups' distributions showed Cooperatives and Private enterprises to be 'better' distributed, reflecting their striving for higher level of efficiency of production. However, the statistical insignificance observed in most cases of Table 5 is likely to be imputed by the relatively low number of observations in some of the groups tested—leading to low power of the test in rejecting the null hypotheses (see SIMAR AND ZELENYUK (2004) for power analysis).

## 5 CONCLUSIONS

In this paper we exercise efficiency analysis specifically for sunflower seeds industry of Ukraine. We also shed some light on the effectiveness of 2000 Land reform in this industry. We exploited the Data Envelopment Analysis to estimate the unknown individual efficiency scores and then used the aggregation theory to obtain estimates of aggregate efficiencies for various types of firms. We also used the statistical bootstrap for making inference on the aggregate efficiencies and for comparison of distributions of individual efficiency scores across these different types of firms.

For the period analyzed, 1998-2002, the DEA approach have shown that sunflower industry in Ukraine had relatively low average efficiency, varying around 50% for virtually all regions analyzed. Analysis of aggregate efficiency scores and of distributions of individual efficiency scores for each region for both models suggests for no significant improvement of efficiency in producing sunflower seeds in both periods, 1998-2000 and 2001-2002.

In order to identify sources of inefficiency each region was explored in more details. For 2001-2002 we selected some most important regions in terms of sunflower seeds production volumes and considered four types/groups of enterprises within each region. These groups were selected according to the current legitimate ownership structures, i.e. Limited Liability Companies (Ltd), Private Companies, Cooperatives, and State-Owned enterprises. Ltd companies proved to be the largest producers of sunflower seeds followed by private companies. Then, these four types/groups of enterprises were also analyzed for group efficiency and distribution of inefficiency, within each region. Such analysis has suggested that, in majority of cases, on the aggregate level, groups performed very similar. But the pattern of distribution showed that Cooperatives and Private enterprises having some tendency to overperform Ltd and State enterprises; but they were as efficient as Ltd firms on the aggregate level. This suggests that at this stage the ownership structure of enterprises does not give convincing arguments in favor of some particular group of firms. These facts witness for no positive development in the operation of industry as a whole; but there were some tendency of some individual enterprises (i.e. Cooperatives and Private) to strive for full efficiency.

## REFERENCES

ASLUND, ANDERS (2002). *Why Has Ukraine Returned to Economic Growth?* Institute for Economic Research and Policy Consulting. Working Paper # 20.

FAO (2002). “Ukraine. The Analysis of Oil-Fat Industry”.

FÄRE R., and ZELENYUK V. (2002). “Input Aggregation and Technical Efficiency”, *Applied Economics Letters*, 9, 635 – 636

FÄRE R., GROSSKOPF S., ZELENYUK V. (2002). “Aggregation Bias and its Bounds in Measuring Technical Efficiency” forthcoming in *Applied Economic Letters.*

FÄRE, R., GROSSKOPF, S. AND C.A.K. LOVELL (1985). The Measurement of Efficiency of Production. Boston, Kluwer-Nijhoff Publishing.

FÄRE R., AND ZELENYUK V. (2003). “On Aggregate Farrell Efficiency Scores,” *European Journal of Operational Research*, 146, 615-20

FARRELL M. J. (1957). “The Measurement of Productive Efficiency”, *Journal of Royal Statistical Society*, Series A, General, 120, part 3, 253-281.

GALUSHKO V., DEMYANENKO S., AND BRUMMER B. (2003). “Farm Efficiency and Productivity Growth in Ukrainian Agriculture”, *Institute for Economic Research and Policy Consulting*. Working Paper No. 20. http:// www.ier.kiev.ua

JOHNSON, S.R., A. BOUZAHER, A. CARRIQUIRY, H. JENSEN, AND P. G. LAKSHMINARAYAN. “Production Efficiency and Agricultural Reform in Ukraine.” *American Journal of Agricultural Economics* 76(1994): 629-635.

KURLAKOVA, LYUBOV A. AND HELEN H. JENSEN. (2002). “Technical Efficiency of Grain Production in Ukraine.” Journal paper No. J-17036. *Center for Agricultural and Rural Development, Iowa State University*

LI, Q. (1996), "Nonparametric Testing of Closeness between Two Unknown Distribution Functions", *Economic Reviews,* 15, pp. 261-274

LI, Q. (1999), "Nonparametric Testing the Similarity of Two Unknown Density Functions: Local Power and Bootstrap Analysis", *Nonparametric Statistics*, 11, pp. 189-213.

MUROVA O., TRUEBLOOD M., AND COBLE K. (2002) “Measurement and Explanation of Technical Efficiency Performance in Ukrainian Agriculture, 1991 - 1996”. *University of Arkansas, Fayetteville, AR 72701*

ROSENBLATT, M. (1956), "Remarks on Some Nonparametric Estimates of a Density Function", *Annals of Mathematical Statistics*, 27, pp. 643-669.

SHEATHER, S. AND M. JONES (1991). "A Reliable Data-Based Bandwidth Selection Mathod for Kernel Density Estimation", *Journal of the Royal Statistical Society,*Series B, 53, 683-90.

SHEPHARD, R.W. (1970). *Theory of Cost Production Functions,* Princeton University Press, Princeton

SILVERMAN, B. W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

SIMAR, L, V. ZELENYUK. (2003), "Statistical Interference for Aggregates of Farrell-type Efficiencies", *Institut de Statistique*, Discussion Paper # 0324, Belgium. http://www.stat.ucl.ac.be

SIMAR, L. AND P.W. WILSON (2000). "A General Methodology for Bootstrapping in Nonparametric Frontier Models", *Journal of Applied Statistique 27, 779-61*.

SIMAR, L. AND ZELENYUK, V. (2004), "On Testing Equality of Distributions of Efficiency Scores Estimated via DEA", Institut de Statistique, Universite Catholique de Louvain, Belgium.

STOSIC, B. AND SAMPAIO DE SOUSA. (2003). "Jackstrapping DEA Scores for Robust Efficiency Measurement". *Anais do XXV Encontro Brasileiro de Econometrica*. *SBE*, 1525-1540.

TACIS (1999), Ukrainian-European Policy and Legal Advice Centre, "Ukrainian Economic Trends"