Technical efficiency of Kansas arable crop farms: a local maximum likelihood approach

Bouali Guesmi*[a], Teresa Serra[a], and Allen Merril Featherstone[b]

[a] Centre de Recerca en Economia i Desenvolupament Agroalimentari (CREDA-UPC-IRTA),
[b] Department of Agricultural Economics, Kansas State University.

*Corresponding author. Parc Mediterrani de la Tecnologia, Edifici ESAB, Avinguda del Canal Olímplic s/n, 08860 Castelldefels, SPAIN. Tel.: +34935521208; Fax: +34935521121. E-mail address: bouali.guesmi@upc.edu

Abstract

The present study uses local maximum likelihood (LML) methods recently proposed by Kumbhakar et al. (2007) to assess the technical efficiency of arable crop Kansas farms. LML techniques overcome the most relevant limitations associated to mainstream parametric stochastic frontier models. Results suggest that Kansas farms reach technical efficiency levels on the order of 90%. These results are compared with another flexible efficiency assessment alternative: the deterministic data envelopment analysis (DEA).

Keywords: Technical efficiency, Nonparametric, Local maximum likelihood approach

JEL classification: C14, Q12, D24

1. Introduction

Technical efficiency is a prerequisite for economic efficiency, which in turn ensures the economic viability and sustainability of a firm. Assessment of firms' technical efficiency levels has drawn broad research interest. Such study is important for producers, as it assists rational input allocation to achieve desired output levels, which strengthens a firms' capacity to face changing market conditions, increasing input costs and economic hardships. It is also relevant for policy makers interested in enhancing firms' economic performance and competitiveness, and promoting economic development.

As is well known, the analysis of technical efficiency assesses to what extent firms are able to maximize their output levels with minimum use of inputs. Two main approaches have been widely used in the efficiency literature namely, parametric (Stochastic Frontier Analysis - SFA) and nonparametric approaches (Data Envelopment Analysis - DEA) (Tzouvelekas et

al., 2001, 2002; Oude Lansink et al., 2002; Sipiläinen and Oude Lansink, 2005; Lohr and Park, 2006). While both encompass several advantages, they are also characterized by some shortcomings. An important difference between these two approaches is that the stochastic production frontier (SPF) allows for the stochastic component of production. This makes this technique suited to assess production processes involving random variables. Most agricultural technologies are stochastic due to unexpected stochastic changes in production (weather influences, for example) and other factors that are not under the control of the farm. Further, agricultural production studies may be affected by measurement and variable omission errors (Coelli, 1995; Chakraborty et al., 2002; Oude Lansink et al., 2002). This makes the SPF approach suited for agricultural performance measurement. The SFA further permits the conduct of conventional statistical tests of hypotheses. However, this approach presents important drawbacks: it relies on the assumption of a parametric functional form representing the production frontier, as well as on a distributional assumption for the random noise and inefficiency error components. Several studies show that technical efficiency results are sensitive to estimation methods and functional form specifications (Ferrier and Lovell, 1990; Coelli and Perelman, 1999; Ruggiero and Vitaliano, 1999; Chakraborty et al., 2001). Inadequate parametric representations of the frontier and the error distributions can lead to biased efficiency estimates (Kumbhakar et al., 2007; Martins-Filho and Yao 2007; Serra and Goodwin, 2009).

Nonparametric DEA techniques overcome the most relevant limitation of SFA: they do not rely on specific functional forms. However, nonparametric approaches do not allow for stochastic variables and measurement errors, which precludes separating inefficiency effects from random noise or random shocks, i.e., all production shortfalls are attributed to the inefficiency term. As a result, technical efficiency ratings obtained from the nonparametric approach (DEA) are generally lower than those obtained under the parametric alternative

(SFA) (Sharma et al., 1999; Puig-Junoy and Argiles, 2000; Wadud and White, 2000). Both methods however have been found to lead to similar rankings of technical performance of decision making units (DMUs).

Recently, a new methodological approach based on local modeling methods has been developed (Kumbhakar et al., 2007) to overcome the limitations of parametric approaches, without foregoing their advantages. In contrast to parametric models, this method does not require strong assumptions regarding the deterministic and stochastic components of the frontier: the parameters characterizing both production and error distribution are allowed to depend on the covariates through a process of localization. As opposed to nonparametric approaches, local modeling methods allow for stochastic variables and variable measurement errors when estimating technical efficiency scores. Furthermore, these techniques accommodate the heterogeneity in the data by deriving observation-specific variances of the inefficiency and noise components of the error term (Serra and Goodwin, 2009). The local modeling approach by Kumbhakar et al. (2007) is based on local maximum likelihood (LML) principles (Fan and Gijbels, 1996).

In spite of the interesting features of this approach, the complexity of implementing the method has limited its use to a few empirical studies. The work by Serra and Goodwin (2009) constitutes a notable exception. Our work contributes to the scarce literature on the use of local modeling techniques to assess technical efficiency. The present study focuses on estimating technical efficiency ratings of a sample of cereals, oilseeds and protein crops (COP) farms in Kansas using flexible LML methods that are compared with the results of DEA techniques. While the existing literature on technical efficiency has broadly compared parametric (SFA) and nonparametric approaches (DEA), to date, there is no study that compares technical efficiency scores obtained under nonparametric (DEA) and LML modeling. In addition, ours constitutes the first study that assesses the efficiency of Kansas

arable crop farms using local modeling approaches (Rowland et al., 1998; Cotton et al., 1999; Serra et al., 2008). The relevance of Kansas as a leading US producer of arable crops makes the analysis especially interesting. In 2010, Kansas generated almost 20% and 50% of total wheat and sorghum produced in the US, respectively. Kansas is also a leading corn and soybean producer, with around 5% of the global US production. The relevant role of Kansas in US arable crop production justifies our decision to study technical efficiency of Kansas arable crop farms.

The paper is organized as follows. In the next section we describe the methodology used in our empirical analysis. The third section presents the data and results from the empirical implementation. We finish the paper with concluding remarks.

2. Methodology

The existing literature that assesses the technical efficiency under which firms operate, has mainly focused on two approaches: the stochastic parametric approach and the deterministic nonparametric method. Several studies have been carried out showing the drawbacks and advantages of each technique. While parametric approaches require strong assumptions regarding specification of the production frontier and the error distribution, that can lead to misspecification issues and biased efficiency estimates, nonparametric approaches do not rely on specific functional forms. However, nonparametric techniques ignore the stochastic component of production, which can also lead to biased technical efficiency measures.

Fan et al. (1996) propose a two-step pseudo-likelihood estimator that does not require specification of the functional form of the production frontier, but still requires assuming a distributional form for the stochastic components of the frontier. A new local modeling parametric approach recently proposed by Kumbhakar et al. (2007) overcomes the limitations

of SFA, without foregoing its advantages. This approach is built upon the LML principle (Fan and Gijbels, 1996) in which the parameters of the stochastic and the deterministic components of the frontier model are localized (flexibilized) with respect to the covariates.

Since our analysis is based on a large number of Kansas farms over a broad geographic region with different climatic conditions, heterogeneity is likely to characterize the sample (different farm sizes, uneven skills, etc). Implementation of the LML approach by Kumbhakar et al. (2007) is suited to deal with heteroscedasticity, as it localizes the standard errors characterizing the distribution of efficiency and noise components of the error term. Based on this approach, we seek to assess the technical efficiency with which COP producing Kansas farms operate and compare efficiency ratings with scores derived by the DEA approach.

The general stochastic frontier models proposed by Aigner et al. (1977) and Meeusen and Van den Broeck (1977) can be specified as follows $Y_i = \beta_0 + \beta^T X_i - u_i + v_i$, where Y_i denotes the observed quantity of output produced by firm $i = 1,...,N$, $X_i \in \Box^d$ is a vector of input quantities required by the production technology, the betas are unknown parameters to be estimated, $u_i > 0$ is the inefficiency term and v_i is a random noise term. The parametric estimation of stochastic frontier models is usually based on maximum likelihood techniques. The joint pdf of (Y, X) is decomposed into a marginal pdf for X, $pdf(x) = p(x)$ and a conditional pdf for Y given x, $pdf(y \mid x) = g(y, \theta(x))$, where $\theta(x) \in \Box^k$ is the localized vector of parameters to be estimated, and g is a function assumed to be known.

Based on the parametric model proposed by Aigner et al. (1977), the conditional pdf for Y given $X = x$ can be specified as: $Y = r(X) - u + v$, where $r(x)$ is the production frontier, $u \mid X = x \Box \left| N\left(0, \sigma_u^2(x)\right) \right|$ and $v \mid X = x \Box N\left(0, \sigma_v^2(x)\right)$, and u and v are assumed to be independently distributed, conditional on X. Following Kumbhakar et al. (2007), the 3-

dimensional local parameter is represented as $\theta(x) = \left(r(x), \sigma_u^2(x), \sigma_v^2(x) \right)^T$ and is approximated using local polynomials. The conditional log-likelihood function $L(\theta) = \sum_{i=1}^{N} \log g\left(Y_i, \theta(X_i) \right)$ is locally approximated by the following mth order local polynomial fit:

$$L_N\left(\theta_0, \theta_1, ..., \theta_m \right) = \sum_{i=1}^{N} q\left(Y_i, \theta_0 + \theta_1 \left(X_i - x \right) + ... + \theta_m \left(X_i - x \right)^m \right) K_H\left(X_i - x \right) \qquad (1)$$

where x represents a fixed interior point in the support of $p(x)$, $q = \log g$, $\theta_j = \left(\theta_{j1}, ..., \theta_{jk} \right)^T$ for $j = 0, 1, ..., m$, and $K_H(u) = |H|^{-1} K\left(H^{-1} u \right)$, where K is a multivariate kernel function and H is assumed to be a positive definite and symmetric bandwidth matrix. The local polynomial estimator is given by $\hat{\theta}(x) = \hat{\theta}_0(x)$ where

$$\left(\hat{\theta}_0(x), ..., \hat{\theta}_m(x) \right) = \arg\max_{\theta_0, ..., \theta_m} L_N\left(\theta_0, \theta_1, ..., \theta_m \right) \qquad (2)$$

Kumbhakar et al. (2007) propose to empirically derive the LML estimator using a local linear fit. By assuming the random noise and inefficiency components to be distributed following a local normal and a half normal distribution, respectively, the conditional probability density function of $\varepsilon = v - u$ is specified as:

$$f\left(\varepsilon \mid X = x \right) = \frac{2}{\sigma(x)} \varphi\left(\frac{\varepsilon}{\sigma(x)} \right) \Phi\left(-\varepsilon \frac{\lambda(x)}{\sigma(x)} \right) \qquad (3)$$

where $\sigma^2(x) = \sigma_u^2(x) + \sigma_v^2(x)$, $\lambda(x) = \sigma_u(x)/\sigma_v(x)$ and $\varphi(.)$ and $\Phi(.)$ represent the probability and the cumulative distribution functions of a standard normal variable, respectively. The local linear parameter is given by $\theta(x) = \left(r(x), \sigma^2(x), \lambda(x)\right)^T$ and the conditional pdf of Y given X is expressed as:

$$g\left(y; \theta(x)\right) = \frac{2}{\sigma(x)} \varphi\left(\frac{y - r(x)}{\sigma(x)}\right) \Phi\left(-\left(y - r(x)\right) \frac{\lambda(x)}{\sigma(x)}\right) \qquad (4)$$

The conditional local log-likelihood function is specified as:

$$L(\theta) \propto \sum_{i=1}^{N} -\frac{1}{2} \log \sigma^2(X_i) - \frac{1}{2} \frac{\left(Y_i - r(X_i)\right)^2}{\sigma^2(X_i)} + \log \Phi\left(-\left(Y_i - r(X_i)\right) \frac{\lambda(X_i)}{\sqrt{\sigma^2(X_i)}}\right) \qquad (5)$$

In the present study, a local linear model for the frontier $r(x_i)$ and a local constant model for the parameters of the error term is used that allows rewriting expression (5) as:

$$L_N(\theta_0, \Theta_1) \propto \sum_{i=1}^{N} -\frac{1}{2} \log \sigma_0^2 - \frac{1}{2} \frac{\left(Y_i - r_0 - r_1^T(X_i - x)\right)^2}{\sigma_0^2} + \log \Phi\left(-\left(Y_i - r_0 - r_1^T(X_i - x)\right) \frac{\lambda_0}{\sqrt{\sigma_0^2}}\right) K_H(X_i - x) \qquad (6)$$

where $\theta_0 = \left(r_0, \sigma_0^2, \lambda_0\right)^T$ and $\Theta_1 = r_1^T$. The local linear estimator of the model is given by $\hat{\theta}_0$:

$$\left(\hat{\theta}_0(x), ..., \hat{\Theta}_1(x)\right) = \arg\max_{\theta_0, \Theta_1} L_N(\theta_0, \Theta_1) \qquad (7)$$

8

Following Jondrow et al. (1982), the efficiency measure for a particular point can be obtained from the following expression:

$$\hat{u}_i = \frac{\hat{\sigma}_0(X_i)\hat{\lambda}_0(X_i)}{1+\hat{\lambda}_0^2(X_i)}\left[\frac{\varphi\left(-\hat{\varepsilon}(X_i)\hat{\lambda}_0(X_i)\big/\hat{\sigma}_0(X_i)\right)}{\Phi\left(-\hat{\varepsilon}(X_i)\hat{\lambda}_0(X_i)\big/\hat{\sigma}_0(X_i)\right)} - \frac{\hat{\varepsilon}(X_i)\hat{\lambda}_0(X_i)}{\hat{\sigma}_0(X_i)}\right] \qquad (8)$$

where $\hat{\varepsilon}(X_i) = Y_i - \hat{r}_0(X_i)$. In the case of variables measured in logs, the efficiency score is given by $\hat{eff}_i = \exp(-\hat{u}_i) \in [0,1]$. Finding a solution to the maximization problem in (7) requires specifying starting values. To do so, we follow Kumbhakar et al. (2007) and start with the local linear least squares estimator of $\hat{r}_0(x)$ and $\hat{r}_1(x)$ and the global ML estimators of $\hat{\sigma}^2$ and λ. The local intercept $\hat{r}_0(x)$ is corrected for the moment condition along the lines of the parametric Modified Ordinary Least Squares (MOLS) estimator. Kumbhakar et al. (2007) recommend using the following expression for such purpose $\hat{r}_0^{MOLS}(x) = \hat{r}_0(x) + \sqrt{2\hat{\sigma}_u^2/\pi}$, where $\hat{\sigma}_u^2 = \hat{\sigma}^2\hat{\lambda}^2\big/(1+\hat{\lambda}^2)$. Hence, initial values for solving (7) are obtained from $\theta_0 = \left(\hat{r}_0^{MOLS}, \hat{\sigma}^2, \hat{\lambda}\right)^T$ and $\Theta_1 = \hat{r}_1(x)^T$.

The product kernel chosen is $h^{-d}\prod_{j=1}^{d} = K\left(h^{-1}(x_j)\right)$, where $K(.)$ represents the Epanechnikov Kernel and d represents the number of covariates. The bandwidth is adjusted for different variable scales and sample sizes and is defined as: $h = h_{base}s_x N^{-1/5}$; where s_x represents the vector of empirical standard deviations of the covariates and N represents the number of observations. The choice of the optimal value for h_{base} is based on the cross validation criterion (CV) proposed by Kumbhakar et al. (2007). The CV, for a given value of h_{base}, is computed by minimizing the following expression:

$$CV\left(h_{base}\right) = \frac{1}{N}\sum_{i=1}^{N}\left[\left(Y_i - \left(\hat{r}_0^{(i)}(x) - u_i^{(i)}\right)\right)\right]^2, \tag{9}$$

where $\hat{r}_0^{(i)}$ and $u_i^{(i)}$ are the leave-one-out versions of the local linear estimators defined above.

As noted above, apart from Kumbakar et al.'s (2007) LML proposal, efficiency of Kansas farms is also assessed by DEA approaches. Following Färe et al. (1994), the DEA linear programming model to assess output-oriented technical efficiency levels can be expressed as:

$$
\begin{aligned}
&\max_{\phi,\lambda} \phi \\
&s.t. \\
&-\phi y_i + Y\lambda \geq 0 \\
&x_i - X\lambda \geq 0 \\
&N1'\lambda = 1 \\
&\lambda \geq 0
\end{aligned}
\tag{10}
$$

where $1 \leq \phi < \infty$, N is the number of farms, X is a $d \times N$ matrix of inputs, Y is a $1 \times N$ matrix of outputs. Technical efficiency scores are given by $1/\phi$. The constraint $N1'\lambda = 1$ is included to allow for variable returns to scale (VRS). As is well known, without such constraint constant returns to scale (CRS) are assumed (Charnes et al., 1994). To test for divergence between the efficiency distributions obtained from LML and DEA methods, the standard Kolmogorov-Smirnov (KS) two-sample (two-tail) test statistic is conducted:

$$D = \max\left|F^a(x,N) - F^b(x,N)\right| \tag{11}$$

where $F^a(x, N)$ represents the empirical distribution function for a sample a with total observations N.

3. Data and results

The empirical application focuses on a sample of Kansas farms that specialize in the production of COP crops. Farm-level data are obtained from farm account records from the Kansas Farm Management Association (KFMA) dataset and cover the period 2000-2010. Data available include farm production and input use, financial and socio-economic characteristics, as well as farm structural characteristics. To ensure that COP is the main farm output, farms whose COP sales represent at least 90% of total farm income were selected. This criterion allows obtaining a relatively homogeneous sample of farms. The dataset is an unbalanced panel that contains 1,258 observations.

We define farm output (y_i) as an implicit quantity index that is computed as the ratio of production in currency units to the output price index. Since information on market prices is unavailable at the farm-level, the Paasche price index is built on the basis of state-level cash unit prices and production data. Output y_i includes the predominant crops in Kansas (Albright, 2002): wheat, corn, soybean and sorghum sales. The inputs considered as explanatory variables are COP land (x_1) measured in acres, total labor input (x_2), mainly composed of family labor, and expressed in annual working units (AWUs), as a fraction of 10-hours per day, chemical inputs (x_3), other inputs (x_4) and capital (x_5). Chemical inputs are defined as a quantity index that includes the use of fertilizers and pesticides, and is obtained by dividing input expenditures by its corresponding price index. Other inputs, also defined as a quantity index, include fuel and seed expenses. Capital input (x_5) aggregates the

value of machinery, other equipment and buildings used in the production process, and is determined by dividing capital value by its corresponding price index. Input prices are measured using national input price indices. Monetary values are measured at constant 2000 prices. Data unavailable from the Kansas database are obtained from the United States Department of Agriculture (USDA) and the National Agricultural Statistics Service (NASS), from which country-level input price indices and state-level output prices and quantities are determined.

Table 1 provides summary statistics for the variables used in the analysis. Sample farms use, on average, 293 AWUs, of which 82% represents unpaid family labor. In contrast to the European Union (EU) arable crop farms that are mainly small holdings with around 116 acres (Farm Accountancy Data Network, FADN 2012), Kansas farms devote 1,278 acres on average to COP production. More than 80% of the COP area is allocated to wheat, soybeans, sorghum and corn production. The average value of farm production (around 154 thousand dollars) almost doubles the EU value (about 84 thousand dollars). However, per acre statistics suggest that EU farms are much more intensive than Kansas farms: while EU farms have an average income of 441 dollars per acre, Kansas income is 122 dollars per acre. Sample farms' investments in machinery and buildings are on the order of 163 thousand dollars. On per acre basis, Kansas farms are less intensive in capital use (150 dollars per acre) relative to the EU with investment ratios on the order of 1,666 dollars per acre (FADN, 2012). To ensure immunity against pests and diseases and to avoid productivity loss due to pest infestations, Kansas farmers spend around 38 thousand dollars annually on chemical inputs. On a per acre basis, expenses in fertilizers and crop protection products are much higher in EU farms (178 dollars per acre versus 29 dollars per acre). Expenses in other inputs, seeds and energy, is rather low compared to chemical input costs and on the order of 24 thousand dollars.

Using the aforementioned variables and following Kumbhakar et al. (2007), we specify the parametric model as a Cobb-Douglas function:

$$\log Y = \beta_0 + \beta_1 \log x_1 + \beta_2 \log x_2 + \beta_3 \log x_3 + \beta_4 \log x_4 + \beta_5 \log x_5 - u + v$$

(12)

It is relevant to note that rigidities associated to this production frontier are overcome by estimating the frontier for each observation in the sample, i.e., flexibility is achieved through varying parameter estimates. To select the bandwidth parameter required to derive the LML estimator of (12), we use the CV procedure described above and evaluated at each sample point. It is worth noting that with multiplicative multivariate kernels, an observation i will only be considered in the LML estimation if all covariates x_i fall into the interval $[x_i - h_i, x_i + h_i)$; where $h_i = h_{base} s_{x_i} N^{-1/5}$. If even one of the components fails to fall into this interval, the observation will not be considered for the estimation. Such procedure requires relatively large values for h_{base} in order to have a sufficiently large subsample of observations to locally estimate the stochastic production frontier. Hence, the more important the sample heterogeneity is, the bigger the required bandwidth. We start with a crude grid of values to then focus on a finer grid for the selection of the optimal h_{base}. Final results show that the bandwidths h_1, h_2, h_3, h_4 and h_5 take values of 2.38, 2.34, 3.37, 3.47 and 2.86, respectively. Once we select the adequate bandwidth for our data, we then derive local parameter estimates.

Descriptive statistics for the variation of the local estimates of σ_u^2 and σ_v^2 are shown in table 2. These statistics confirm the presence of heteroscedasticity and indicate an important degree of variation among observations regarding the shares of the inefficiency term to the noise term ($\lambda = \sigma_u^2 / \sigma_v^2$).

Figure 1 illustrates the variation of the parameters of the deterministic component of the frontier. Since we use a Cobb–Douglas functional form for our model, the coefficients represent input elasticities. The variation of the localized estimates suggests that assuming the same input elasticities for all observations may not be reliable. Variation is specially relevant for land, with an elasticity that ranges from 25% to 45%, followed by chemical inputs and capital, that have an elasticity fluctuating from 26% to 38% and 18% to 30%, respectively. Input elasticities indicate that farms operate under constant returns to scale with a mean scale elasticity equal to 0.997 and a standard deviation of 0.045 (table 3).

As expected, localized elasticity estimates are positive for a majority of farms in that none of the inputs considered is over-utilized. Noteworthy is the fact that, for some observations, the labor elasticity is negative. This is not surprising given the share of family labor in our sample farms. Since this labor type usually involves an opportunity cost but not a direct cost, incentives to use it efficiently may be less strong than for other inputs.

Production elasticity estimates indicate, on average, that an increase in chemical input use has the highest potential to increase output, followed by land, capital, other inputs and labor (table 3). The low contribution of labor to farm productivity in Kansas farms can once more be attributed to the high share of family labor. The fact that capital, land and other inputs have lower elasticities than chemicals inputs suggests that the latter are used less intensively.

Table 4 illustrates the distribution of the localized efficiency estimates. Results show a high average technical efficiency score, on the order of 0.90, indicating that farmers reach 90% of their maximum potential output. Therefore, our sample farms could increase their output by 10% by efficiently using their inputs. Our results differ from those in Serra et al. (2008) who used the same database, but focused on the period 1998-2001. Through Kumbhakar's stochastic frontier model (2002), Serra et al. (2008) obtained mean technical

inefficiency levels of 0.30, versus 0.10 in our analysis. The use of different methodologies or farmers' performance improvement over time can explain differences in efficiency scores across studies. However, our results are closer to other findings by Rowland et al. (1998) for a sample of Kansas swine operations from 1992 through 1994, or Cotton et al. (1999), for a sample of multi-output Kansas farms during the period 1985 to 1994. Both authors used nonparametric DEA techniques to derive efficiency estimates and obtained mean efficiency scores of 0.89 and 0.91, respectively.

High technical efficiency levels are associated with low production costs and higher chances of firm's economic viability. Technical efficiencies range from a minimum of 0.19 to a maximum of 0.99 indicating important dispersion and heterogeneity within Kansas farms. Almost one half of the observations display high performance levels presenting efficiency ratings greater than 0.90.

DEA efficiency scores under CRS (0.81) assumption are, on average, lower than those derived from the LML approach. Under VRS, however, efficiency ratings are much closer to LML results (0.92).[1] At the 5% level of significance, the KS[2] test indicates that the difference in efficiency scores derived from DEA and LML techniques is statistically significant (table 5). Given the fact that LML is a technique that overcomes the most relevant limitations of DEA methods, its reliability may be higher.

---

[1] DEA results suggest that Kansas farms do not operate at optimal scales.

[2] The nonparametric Li (1999) test has been also computed indicating that the two distribution obtained from LML and DEA methods are equal and the null hypothesis cannot be rejected at the 5% level of significance with p-value 0.35 and 0.33 for LML vs. VRS and LML vs. CRS respectively. However, the power of this test is very sensitive to the dimensionality and sample size which can conduct to misleading impression of equality of distributions (Simar and Zelenyuk, 2006).

4. Concluding remarks

The relevance of deriving reliable technical efficiency scores to assist firms' management decisions as well as policy design, makes it essential to use methodologies that produce farm-level non-biased efficiency ratings. The parametric SFA and the nonparametric DEA approach have focused the attention of mainstream efficiency literature. Both approaches have been widely criticized for their shortcomings that may lead to biased efficiency estimates.

Recently, Kumbhakar et al. (2007) proposed a new approach, namely the LML method. The method estimates the parameters of the deterministic and stochastic components of the frontier locally. LML methods overcome the shortcomings of SFA without foregoing their advantages. LML techniques are used in this article to assess the efficiency levels achieved by Kansas farms specialized in cereals, oilseeds and protein crops (COP) production and compares them with those obtained from flexible DEA models. Farm-level data obtained from farm account records from the KFMA dataset covering the period 2000-2010 are used.

Empirical results support the relevance of using the LML approach through the variation in the localized parameter estimates, representing the variance of the composite error term and input elasticities. Results show high mean efficiency scores (0.90) indicating that farmers could increase their output by 10% keeping their input bundle constant. Technical efficiency scores derived from the LML approach are higher (lower) than those of the DEA model under CRS (VRS). According to KS test, the efficiency scores obtained from DEA and LML have different distributions and the difference is statistically significant. Since LML allow for both stochastic error terms, as well as for flexibility in the functional form representing the frontier function, we suggest that efficiency scores derived under LML

maybe more reliable and less biased than efficiency ratings under nonparametric DEA alternatives.

Our research can be extended in many different ways. Different methodological innovations to assess efficiency have been recently introduced in the literature. Noteworthy are the refinements regarding the measurement of technical efficiency in the presence of uncertainty through state-contingent techniques (Chambers and Quiggin, 2000). Failure to properly allow for risk can lead to biased efficiency estimates (O'Donnell et al., 2010). Other innovations in the technical efficiency literature include dynamic efficiency measurement that does not rely on the assumption of firm's ability to adjust instantaneously and that allows for the dynamic linkages of production decisions (Serra et al., 2011). Extension of LML methods to a consideration dynamic issues constitutes another area that merits further attention.

References

Aigner, D., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production functions models. J. Econometrics 6, 21-37.

Albright, M. L., 2002. Kansas Agriculture: An economic overview for 2001, Research Papers and Presentations, Kansas State University (KSU) Research and Extension. Accessed November 2011, available at http://www.agmanager.info/farmmgt/income/papers/.

Chakraborty, K., Biswas, B., Lewis, W.C., 2001. Measurment of technical efficiency in public education: A stochastic and non-stochastic production function approach. Southern Econ. J. 67, 889-905.

Chakraborty, k., Misra, S., Johnson P., 2002. Cotton farmers' technical efficiency: Stochastic and non-stochastic production function approaches. Agr. Resource Econ. Rev. 31, 211-220.

Chambers, R., Quiggin, J., 2000. *Uncertainty, production choice and agency: The state-contingent approach*. Cambridge University Press, Cambridge, UK.

Charnes, A., Cooper, W., Lewin, A., Seiford, L., 1994. *Data envelopment analysis, theory, methodology and applications*. Kluwer Academic Publishers, Norwell, MA.

Coelli, T.J., 1995. Recent developments in frontier modeling and efficiency measurement. Australian J. Agr. Resource Econ. 39, 219-245.

Coelli, T.J., Perelman, S., 1999. A comparison of parametric and non-parametric distance functions: With application to European railways. Eur. J. Oper. Res. 117, 326-339.

Cotton, M.K., Langemeier, M.R., Featherstone, A.M., 1999. Effects of weather on Multi-Output: Efficiency of Kansas Farms. Journal of the ASFMRA 62, 85-91.

FADN, 2012. FADN public database. Accessed April 2012, available at http://ec.europa.eu/agriculture/rica/database/database_en.cfm.

Fan, J., Gijbels, I., 1996. *Local polynomial modeling and its applications*. Chapman and Hall, London, UK.

Fan, Y., Li, Q., Weersink, A., 1996. Semiparametric estimation of stochastic production frontier models. J. Bus. Econ. Statist. 14, 460-468.

Färe, R., Grosskopf, S., Lovell, C.A.K., 1994. *Production frontiers*. Cambridge University Press, Cambridge, UK.

Ferrier, R.S., Lovell, C.A.K., 1990. Measuring cost efficiency in banking: An econometric and linear programming evidence. J. Econometrics 46, 229-245.

Jondrow, J., Lovell, C.A.K., Materov, I.S. and Schmidt, P. (1982). On the estimation of technical inefficiency in stochastic frontier production models. J. Econometrics 19, 233-238.

Kumbhakar, S.C., 2002. Specification and estimation of production risk, risk preferences and technical efficiency. Amer. J. Agr. Econ. 84, 8-22.

Kumbhakar, S.C., Byeong, U.P., Simar, L., Tsionas E.G., 2007. Nonparametric stochastic frontiers: A local maximum likelihood approach. J. Econometrics 137, 1-27.

Li, Q., 1999. Nonparametric testing the similarity of two unknown density functions: Local power and bootstrap analysis. J. Nomparametr. Stat. 11, 189-213.

Lohr, L., Park, T.A., 2006. Technical efficiency of US organic farmers: The complementary roles of soil management techniques and farm experience. Agr. Resource Econ. Rev. 35, 327-338.

Martins-Filho, C., Yao, F., 2007. Nonparametric frontier estimation via local linear regression. J. Econometrics 141, 283-319.

Meeusen, W., Van Den Broek, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. Int. Econ. Rev. 18, 435-444.

O'Donnell, C.J., Chambers, R.G., Quiggin, J., 2010. Efficiency analysis in the presence of uncertainty. J. Productiv. Anal. 33, 1-17.

Oude Lansink, A., Pietola, K., Bäckman, S., 2002. Efficiency and productivity of conventional and organic farms in Finland 1994–1997. Europ. Rev. Agr. Econ. 29, 51-65.

Puig-Junoy, J., Argiles, J.M., 2000. Measuring and explaining farm inefficiency in a panel data set of mixed Farms. Pompeu Fabra University, Working Paper, Barcelona, Spain, July 25.

Rowland, W.W., Langemeier, M.R., Schurle, B.W., Featherstone, A.M., 1998. A nonparametric efficiency analysis for a sample of Kansas swine operations. J. Agr. Appl. Econ. 30, 189-199.

Ruggiero, J.,Vitaliano, D., 1999. Assessing the efficiency of public schools using data envelopment analysis and frontier regression. Contemporary Econ. Pol. 17, 321-331.

Serra, T., Goodwin, B.K., 2009. The efficiency of Spanish arable crop organic farms, a local maximum likelihood approach. J. Productiv. Anal. 3, 113-124.

Serra, T., Oude Lansink, A., Stefanou, S., 2011. Measurement of dynamic efficiency: A directional distance function parametric approach. Amer. J. Agr. Econ. 93: 756-767.

Serra, T., Zilberman, D., Gil J.M., 2008. Farms' technical inefficiencies in the presence of government programs. Australian J. Agr. Resource Econ. 52, 57-76.

Sharma, K.R., Leung, P., Zaleski, H.M., 1999. Technical, allocative and economic efficiencies in swine production in Hawaii: A comparison of parametric and nonparametric approaches. Agr. Econ. 20, 23-35.

Simar, L., Zelenyuk, V., 2006. On testing equality of distributions of technical efficiency scores. Econometric Rev. 25, 497-522.

Sipiläinen, T., Oude Lansink A., 2005. Learning in organic farming-an application on Finnish dairy farms. Manuscript presented at the XIth Congress of the EAAE. Copenhagen, Denmark, August 24–17.

Tzouvelekas, V., Pantzios, C.J., Fotopoulos, C., 2001. Technical efficiency of alternative farming systems: The case of Greek organic and conventional olive-growing farms. Food Pol. 26, 549–569.

Tzouvelekas, V., Pantzios, C.J., Fotopoulos, C., 2002. Empirical evidence of technical efficiency levels in Greek organic and conventional farms. Agr. Econ. Rev. 3, 49-60.

USDA, 2011. Agricultural prices summary, crop values annual summary and crop production annual summary. Accessed August 2011, available at http://usda.mannlib.cornell.edu/MannUsda.

Wadud, A., White, B., 2000. Farm household efficiency in Bangladesh: A comparison of stochastic frontier and DEA methods. Appl. Econ. 32, 165-73.

Table 1. Summary statistics for the variables of interest

Variable	Mean	Standard deviation
Total output (index)	154,193.14	164,521.51
Capital (index)	162,547.25	158,754.89
Land (acres)	1,277.89	1,103.34
Labor (AWU)	292.68	252.84
Chemicals inputs (index)	38,296.45	41,985.78
Other inputs (index)	24,398.16	25,388.22
Statistics on a per acre basis		
Total output (dollars/acre)	122.50	66.52
Capital (dollars/acre)	150.36	131.49
Labor (AWU/acre)	0.24	0.14
Chemicals inputs (dollars/acre)	29.11	16.98
Other inputs (dollars/acre)	19.48	12.68

Table 2. Summary statistics for the local estimates of σ_u^2, σ_v^2 and λ

	σ_u^2	σ_v^2	λ
Maximum (100%)	1.14	0.26	22.20
Third quartile (75%)	0.03	0.10	0.59
Median (50%)	0.02	0.09	0.48
First quartile (25%)	1.93E-5	0.09	0.01
Minimum (0%)	6.93E-7	7.59E-4	0.30E-2

Table 3. Distribution of production and scale elasticities for Kansas Farms

Elasticities with respect to	Estimate	Standard deviation
Land area	0.250	0.109
Labor	0.036	0.075
Capital	0.217	0.058
Chemical inputs	0.295	0.054
Other inputs	0.199	0.062
Returns to scale	0.997	0.045

Table 4. Frequency distribution of technical efficiency scores

TE: Range (%)	Observations			(%)		
	LML[1]	VRS[2]	CRS[3]	LML	VRS	CRS
<80	78	2	645	6.20	0.16	51.27
80-85	130	31	449	10.33	2.46	35.69
85-90	435	363	103	34.58	28.86	8.19
90-95	244	657	31	19.40	52.23	2.46
95-100	371	205	30	29.49	16.30	2.38
Mean	0.90	0.92	0.81			
Standard deviation	0.08	0.03	0.05			
Minimum	0.19	0.78	0.68			
Maximum	0.99	1.00	1.00			

[1]LML: local maximum likelihood. [2]VRS: variable returns to scale. [3]CRS: constant return to scale.

Table 5. Kolmogorov-Smirnov test

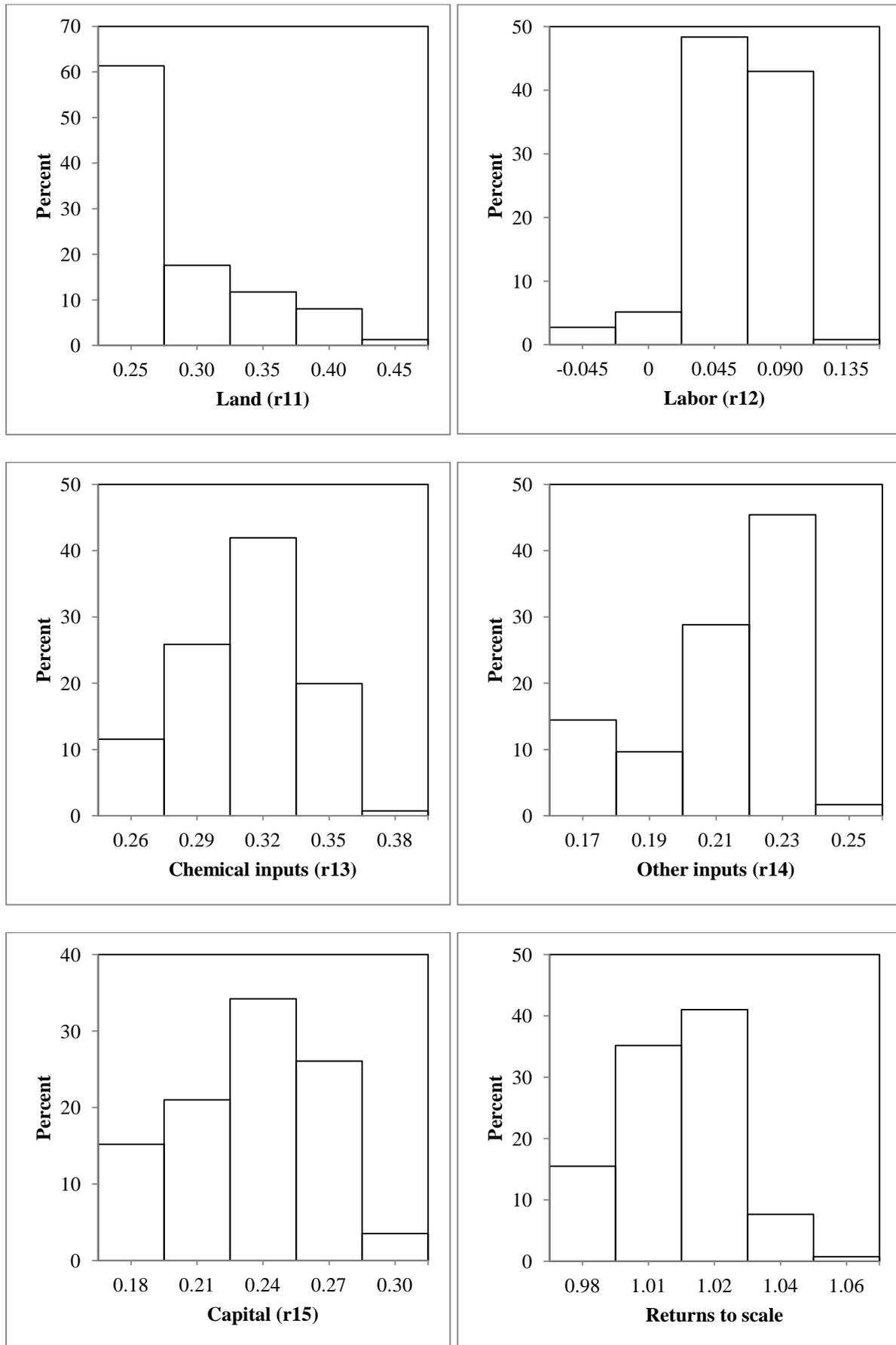| Test	Value	p-value
LML vs. DEA VRS	0.257	0.000
LML vs. DEA CRS	0.711	0.000

Fig. 1 Distribution of localized estimates of input elasticities and returns to scale