Comparative Analysis of Stochastic Frontier Partially non-parametric and Stochastic Frontier Parametric Methods Case Study: Measuring Cost Efficiency in Wheat Production in Iran

Ebrahim Moradi 1*, Mosayeb Pahlavani 2, Ahmad Akbari 3 and Hossain Mehrabi Bashrabadi 4

Received: 1 March 2013, Accepted: 18 May 2013

Abstract

Most research on performance evaluation of agricultural crops in Iran analyzes technical efficiency with estimation of stochastic production function. In this paper in spite of past research we survey cost efficiency with panel data. Information about input prices, yield and production cost per hectare collected for 28 provinces in 10 years and stochastic frontier cost function estimated with panel data in two methods, parametric stochastic frontier and partially non-parametric stochastic frontier. The results show that according to parameter significance and discretion of production structure parametric methodology is more suitable than non-parametric methodology. Land rent (price) had maximum influence and chemical fertilizer price had minimum influence on frontier production cost per hectare. Cost efficiency of wheat production in Iran is suitable and over 90 percent. Surveyed period Khuzestan province had maximum cost efficiency and Yazd province had minimum cost efficiency in wheat production. In whatever province where efficiency is low, agricultural education and knowledge needed to improve further.

Keywords: Cost Efficiency, Partially Non-Parametric Method, Stochastic Frontier Method, Iran

1 Ph.D. Student, University of Sistan and Baluchestan, Iran.
2 Associate Professor, University of Sistan and Baluchestan, Iran.
3 Professor, University of Sistan and Baluchestan, Iran.
4 Professor, University of Bahonar Kerman, Iran.
* Corresponding author’s email: ebmoradi31@gmail.com
INTRODUCTION
Today, grain has unique and important role in consuming pattern at any country. Wheat, barley, rice and maize have cultivation at highest level in Iran as Wheat 73.17%, barley 18.45%, rice 5.90%, and maize 2.48% of the cultivation has devoted to themselves (Agricultural Statistics, Iran, 2010). Wheat planted area was estimated 6.65 million hectares in 2010, that 36.75 percent is produced by dry farming and 63.25 percent is irrigated farming. Khorasan razavi with 9.21 percent of total wheat of Iran lands allocated highest level planted wheat area to himself. After this province, Kordestan, Fars, Hamadan, eastern Azarbaygan, Zangan and Kermanshah with 8.27%,6.91%,6.75%,6.6% and 6.40 %, allocated second rank to seventh rank to themselves. Therefore, more than half (50.76%) of wheat lands picked up in these seven province. Despite that Fars province has Third place in planted area, it has been in first place with 36.1% of total wheat production and Khuzstan, Khorasan Razavi, Golestan, Kermanshah, Hamadan and Eastern Azarbaygan with 8.75%, 8.36%, 8.17%, 6.11%, 5.69% and 5.43 %, allocated second rank to seventh rank to themselves. (Costs of agricultural production, 2011).

Most of study in Iran in concerned with measuring the performance of agricultural crops production analyzed estimation of stochastic frontier production function and technical efficiency. Since the accuracy of input prices data is higher than amount of inputs in agricultural sector macro data, cost function is suitable in efficiency study. Comparison of the results of the parametric and non parametric methods can show that whether application of nonparametric method Compared to the parametric method is an advantage or not.

Mosavi and Khalilian (2005), with purpose of estimating technical efficiency of some wheat farm in Shahr Kord city, used Translog stochastic frontier production function and estimated model with panel data. whose results show that average technical efficiency is 78%. Gahani and Asghari (2006), surveyed mathematical structure of production wheat cost function in framework of panel data and non frontier model where they used non frontier Translog cost function with cost share equations. After model estimation, they analyzed substitute and complement relation among inputs. Results from this study shows that fertilizer and seed are complement, and machinery and labor are complement.

Zeram Nejad and Yousefi Hagiaabad (2008), for survey technical efficiency of wheat production at different province of Iran have used DEA and stochastic frontier production function methods. They have estimated technical efficiency by using panel data and obtain that the mean of wheat production in Iran is 57% by parametric method as well as the mean of wheat production is 84% by non-parametric method. Moradi Shahrbabak (2008), for estimating the wheat production efficiency at Baft county of Kerman province preferred using cob-Douglass stochastic frontier production function than by duality principle was derivated cost function and got a result that average efficiency of technical, allocative and economic are 88%,84% and 74% .Respectively. The method used for decomposition of efficiency in his paper is not suitable and compatible. Haghiri (2003) has used stochastic non-parametric frontier production function and parametric production function for measuring technical efficiency of north American dairy industry. His motivation for using stochastic non-parametric frontier estimates comes from the fact that there are problems inherent in the structure of stochastic parametric frontier models. The results show that the overall mean of technical efficiency obtained from translog function for all regions is higher than that of the corresponding values obtained from the non-parametric approaches. Both parametric and nonparametric methodologies indicated evidence of differences between the mean technical efficiency of dairy farms in all regions. Cesar Revoredo, et al. (2009) estimate indicators of farm efficiency for the period 1989 to 2008 by farm type and to analyses they used stochastic frontier cost function. The results indicate while mixed farms and lowland farms have maintain their levels of efficiency. LFAfarms have seen their efficiency reduced since approximately 2004 or 2005 (especially LFAsheep farm spe-
cialists). Also, the analysis shows that there seems to be an increase in the dispersion of farmers in terms of efficiency for some farm types in periods of change in agricultural policy.

Considering uncontrollable factors in agriculture, the DEA approach, because of the disregard random effects, has a higher error than stochastic frontier approach. Panel data estimation has higher accuracy than cross-sectional data. According to data quality in agricultural sector of Iran, using of cost function can provide better results.

Despite the fact that most past study in Iran surveyed technical efficiency by estimating production function, this study pay attention to frontier cost function and so stochastic frontier partially non-parametric is introduced. Next, the theoretical basis of the parametric and non-parametric methods of calculating cost efficiency was surveyed and Information about prices, inputs, yield and cost of production per hectare in a 10 year period has been collected for 28 provinces. Stochastic frontier cost function with panel data was estimated by both parametric and nonparametric methods and Analysis and comparison of the results of the estimates are analyzed and compared.

MATERIALS AND METHODS

We assume that we have observation a panel of producers through time periods. Also we assume that the deterministic kernel of the stochastic cost frontier takes cob-douglas form and initially cost efficiency is time invariant. Thus cost frontier model as

$$\ln E_{it} = \beta_0 + \beta_y \ln y_{it} + \beta_n \ln w_{nit} + v_{it} + u_{it}$$  \hspace{1cm} (1)

Where $v_{it}$ represents error term, represents time invariant cost efficiency, and $\sum \beta_n = 1$ ensures homogeneity of degree 1 of cost frontier in input prices. We make the following assumption on the error components in the stochastic cost frontier model.

i. $v_{it} \sim iid \ N(0, \sigma^2_v)$

ii. $u_{it} \sim iid \ N^+(0, \sigma^2_u)$

iii. $u_{it}$ and $v_{it}$ are distributionally independent of each other, and of regressors.

The log likelihood function for a sample of producers, each observed for periods of time, becomes

$$\ln L = \text{constant} - \frac{I(T-1)}{2} \ln \sigma^2_v - \frac{I}{2} \ln(\sigma^2_v + T\sigma^2_u) + \sum \ln \left[1 - \Phi \left(-\frac{\mu_v}{\sigma_v}\right) \right] \left(\frac{\varepsilon_{it}}{2\sigma_v^2} \right)^2 - \frac{1}{2} \sum \left(\frac{\mu_v}{\sigma_v}\right)^2 \right)$$  \hspace{1cm} (2)

Where and $\mu_v = T\sigma_v^2 \varepsilon / (\sigma_v^2 + T\sigma_u^2)$ and $\sigma_v^2 = \sigma_v^2 / (\sigma_v^2 + T\sigma_u^2)$. This log likelihood function can be maximized with respect to the parameters. To obtain maximum likelihood estimates of $\beta$, $\sigma_v^2$, and $\sigma_u^2$.

The conditional distribution of $(u|\varepsilon)$ is

$$f(u|\varepsilon) = \frac{1}{(2\pi)^{\frac{1}{2}} \sigma_u} \exp \left(\frac{\varepsilon^2}{2\sigma_u^2}\right)$$  \hspace{1cm} (3)

Which is the density function of a variable distributed as $N^+(\mu, \sigma^2)$. Either the mean or the mode of this distributions can be used as a point estimator of cost efficiency, and we have

$$E(u_i|\varepsilon_i) = \mu_v + \sigma_u \left[\frac{\phi(-\mu_v/\sigma_v)}{1 - \Phi(-\mu_v/\sigma_v)} \right]$$  \hspace{1cm} (4)

And

$$M(u_i|\varepsilon_i) = \begin{cases} 
\mu_v & \text{if } \varepsilon_i \geq 0 \\
0 & \text{otherwise} 
\end{cases}$$  \hspace{1cm} (5)

Either can be substituted into to $CE_i = \exp \{-u_i\}$ obtain producer-specific estimate of time-invariant cost efficiency. An alternative estimator is provided by the minimum squared error predictor

$$E(\exp \{-u_i\}|\varepsilon_i) = \frac{1 - \Phi(\sigma_v(-\mu_v/\sigma_v))}{1 - \Phi(-\mu_v/\sigma_v)} \exp \left(-\mu_v + \frac{1}{2}\sigma_v^2\right)$$  \hspace{1cm} (6)

Confidence intervals for any of these estimators can be calculated (Kumbhakar and Lovell, 2000).

Most methods of measuring efficiency based on the MLE estimation such as in “Pooled data model”, “Battese and Coelli model”, “pitt and lee model” and “True random effects Green model” In addition, “Cornwell, Schimidt, and Sickles model” based on OLS estimation and non frontier function (Green, 2011). In this study we used “pitt and lee model” this model is similar to “Battese and Coelli model” With the difference that one efficiency score in time period for every firm is estimated, therefore
this model is faced with less error. Another method that can be used to measurement cost efficiency is non-parametric method. The locally weighted scatter plot smoothing, or the locally regression model, is introduced by Cleveland (1979) and developed by Fan (1992; 1993) and Hastie and Loader (1993). The basic idea of using the LOWESS method is to find a point in the space of the predictors and then search the neighborhood points that are smoothed using surface smoothers to estimate the mean response function. For instance, we may consider any point \( x \), a so-called local observation, in the space of the predictors. Estimating a local regression model can be specified through different approaches. In a local regression we attempt to find a neighborhood containing the initial point \( x \) in which the regression surface is well approximated by a function from a specific parametric point of view. Therefore, our specification from the local regression model leads to methods of fitting the response function. The method consists of smoothing the response as a function of the predictors. Our goal is to smooth \( s(x_0) \) in which \( s(.) \) depicts the scatter plot smoother functions, using nearest neighborhoods. First, we identify the nearest neighbors of \( x_0 \) which are denoted by \( \Omega(x_0) \). In the second step, we determine the furthest near-neighbor observation from \( x_0 \) and compute the distance between these two points. That is, we calculate \( \Delta(x_0) = \max_{\Omega(x_0)} |x_0-x| \) next, using the tri-cube weight function (Cleveland et al., 1993), we assign weights to each point in \( \xi_I \) as

\[
\xi(u) = (1-u^3), \text{ for } 0 \leq u \leq 1
\]

In the last step, and by using the weights computed in the third step, we obtain the Scatter plot smoother \( s(x_0) \) value at the initial point by applying the weighted least-squares (WLS) of response variable \( y \) to predictors \( x \) in the domain of \( \Omega(x_0) \). Locally weighted scatter plot smoothers are popular among statisticians and econometricians for at least two reasons. First, these methods are able to produce robust results in respect to the outliers. Second, with the scatter plot smoothers, we are able to easily find the neighborhoods for the target point. Econometricians prefer using the former method since it is less biased, as compared to the latter method (Haghiri, 2003).

**RESULTS AND DISCUSSION**

Data for this study is collected from agricultural crops cost production bank and statistical year-books of agricultural ministry of Iran. Data related to every province from 2000 to 2010 was collected for 28 provinces at 10 years time period. The cost efficiency of wheat production was modeled in a six input and single output framework. Quantity of wheat produced per hectare at any province was used to measure output. Price of Six inputs were used for cost efficiency analyses which are fertilizer, seed,
agricultural pesticides, labor and water prices. In order to aggregate fertilizer and agricultural pesticides input prices we used the weighted average. Descriptive statistics of input price and output is given in table 1.

We were encountered with limitations in machinery and water price. The water price wasn’t available; therefore we used water cost per hectare instead of water price. Unfortunately, suitable variable wasn’t available for machinery price.

Cost efficiency was calculated by two methods namely parametric and partially non-parametric. And we had compared results from two methods. At first model was specified in translog form, unfortunately, we encountered with singular error, after elimination of this error and estimation model and were not significant. Therefore we used cob-Douglas specification form.

\[ \ln E_{it} = \beta_0 + \beta_1 \ln y_{it} + \sum_n \beta_n \ln w_{n,it} + v_{it} + u_{it} \quad (9) \]

The cost function is homogenous degree one at input price. Therefore \( c(y_{it}, \lambda w_{nit}; \beta) = \lambda c(y_{it}, \lambda w_{nit}; \beta) \) and \( \lambda > 0 \). We should impose \( \beta_k = 1 - \sum_{n\neq k} \beta_n \) or we can change equation (9)

\[ \ln \left( \frac{E_{nit}}{w_{nit}} \right) = \beta_0 + \beta_1 \ln y_{it} + \sum_n \beta_n \ln \left( \frac{w_{n,it}}{w_{nit}} \right) + v_{it} + u_{it} \quad (10) \]

We used seed price for normalization variables and after model estimation with pitt and lee (1981) and LOWESS methods by Nlogit 5.0 software package and calculated cost efficiency of the type

\[ CE_{it} = \exp E \{ -uit \} \]

A likelihood ratio test against the hypothesis of no inefficiency follows the variance estimates. Appropriate tables for the mixed chi squared test used here are given in Kodde and Palm (1986). LR test for inefficiency vs. OLS only is

\[ \text{Chi – sq} = 2 \times [\log I(s\bar{f}) - \log I(\bar{s}l)] \quad (12) \]

Table 2 shows that there is not inefficiency in wheat production hypothesis that cannot be accepted and we should estimate model in frontier manner.

The stochastic frontier cost function estimation results was showed in table 3. In cost function estimation based on pooled data and panel data only normalized price of agricultural pesticides hasn’t significant effect on Production cost per hectare wheat and other coefficient are significant at 99%. Normalized land price (land rent) has maximum effect on Production cost per hectare wheat, as one percent increase at land price 0.24 percent increase cost production and normalized fertilizer price has minimum effect on production cost as one percent increase at land price lead to only 0.14 percent increase cost production.

The results of Cost function estimation based on panel data show that log likelihood is increased

<table>
<thead>
<tr>
<th>Table 2: LR test for inefficiency vs. OLS only</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>statistics</strong></td>
</tr>
<tr>
<td>Chi-sq</td>
</tr>
</tbody>
</table>

Kodde-Palm C*: 95%: 2.706, 99%: 5.412

<table>
<thead>
<tr>
<th>Table 3: Models estimation results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>variables</strong></td>
</tr>
<tr>
<td>constant</td>
</tr>
<tr>
<td>( \ln y_{it} )</td>
</tr>
<tr>
<td>( \ln (w_{nit}/w_{nit}) )</td>
</tr>
<tr>
<td>( \ln (w_{nit}/w_{nit}) )</td>
</tr>
<tr>
<td>( \ln (w_{nit}/w_{nit}) )</td>
</tr>
<tr>
<td>( \ln (w_{nit}/w_{nit}) )</td>
</tr>
<tr>
<td>( \ln (w_{nit}/w_{nit}) )</td>
</tr>
<tr>
<td>( \lambda = \sigma/u_0 )</td>
</tr>
<tr>
<td>( \lambda = (\sigma_0^2 + \sigma_0^2)^{1/2} )</td>
</tr>
<tr>
<td>( \sigma_0 )</td>
</tr>
<tr>
<td>( \sigma_0 )</td>
</tr>
<tr>
<td>Log likelihood</td>
</tr>
</tbody>
</table>

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.
significantly, only normalized price of agricultural pesticides hasn’t significant effect on Production cost per hectare wheat and other coefficients are significant at 99%. Normalized land price (land rent) has maximum effect on Production cost per hectare wheat, as one percent increase at land price lead to 0.24 percent increasing cost production similar to pooled data model.

The results of Cost function estimation based on LOWESS method is to some extent different from panel data method. Yield per hectare and agricultural pesticides price haven’t significant effect on production cost while other coefficient are significant at 95%. Land price have maximum effect on production cost and fertilizer price have minimum effect and this is similar to panel data method.

After cost function estimation and calculation cost efficiency, provinces ranked based on cost efficiency calculated by panel data and LOWESS methods that present in table 4.

Khuzestan and Bousher had maximum cost efficiency in wheat production, and Yazd province had minimum cost efficiency in wheat production at panel data method (Table 5).

Mean of efficiency in parametric method is 90% and in non-parametric method (LOWESS) is 97%. Standard deviation in nonparametric method is low and efficiency cost deference among provinces is inconsiderable as in this method minimum cost efficiency is 97% and maximum cost efficiency is 99%.

The kernel density estimator is a device used to describe the distribution of a variable non-parametrically, that is, without any assumption of the underlying distribution (silverman, 1986). The kernel density function for a single variable is computed using

$$f(x) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{x-X_i}{h}\right)$$

(13)

We plotted kernel density for EUP and EUNP that is displayed in figure 1.

Figure 1 shows Difference between two methods. EUP distribution is more suitable than EUNP. Although in cross section data results from parametric and non-parametric method is more similar, in this study that we used panel data, the results of parametric method is more suitable than non-parametric method.

### Table 4: Cost efficiency and province rank

<table>
<thead>
<tr>
<th>Province</th>
<th>Cost efficiency Parametric method</th>
<th>Rank</th>
<th>Cost efficiency LOWESS method</th>
<th>Rank</th>
<th>Province</th>
<th>Cost efficiency Parametric method</th>
<th>Rank</th>
<th>Cost efficiency LOWESS method</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khuzestan</td>
<td>0.995</td>
<td>1</td>
<td>0.986</td>
<td>1</td>
<td>Mazandaran</td>
<td>0.907</td>
<td>15</td>
<td>0.975</td>
<td>20</td>
</tr>
<tr>
<td>Bushier</td>
<td>0.989</td>
<td>2</td>
<td>0.985</td>
<td>2</td>
<td>Golesan</td>
<td>0.900</td>
<td>16</td>
<td>0.976</td>
<td>16</td>
</tr>
<tr>
<td>Kurdistan</td>
<td>0.955</td>
<td>3</td>
<td>0.979</td>
<td>4</td>
<td>Azerbaijan sharghi</td>
<td>0.898</td>
<td>17</td>
<td>0.974</td>
<td>23</td>
</tr>
<tr>
<td>Elam</td>
<td>0.950</td>
<td>4</td>
<td>0.978</td>
<td>10</td>
<td>Azerbaijan gharbi</td>
<td>0.896</td>
<td>18</td>
<td>0.973</td>
<td>24</td>
</tr>
<tr>
<td>Kerman</td>
<td>0.940</td>
<td>5</td>
<td>0.979</td>
<td>8</td>
<td>Kermanshah</td>
<td>0.893</td>
<td>19</td>
<td>0.977</td>
<td>13</td>
</tr>
<tr>
<td>Lorstan</td>
<td>0.929</td>
<td>6</td>
<td>0.978</td>
<td>9</td>
<td>Khorasan</td>
<td>0.886</td>
<td>20</td>
<td>0.979</td>
<td>7</td>
</tr>
<tr>
<td>Tehran</td>
<td>0.929</td>
<td>7</td>
<td>0.979</td>
<td>5</td>
<td>Qom</td>
<td>0.882</td>
<td>21</td>
<td>0.975</td>
<td>19</td>
</tr>
<tr>
<td>Hormozgan</td>
<td>0.927</td>
<td>8</td>
<td>0.979</td>
<td>6</td>
<td>Ardabil</td>
<td>0.881</td>
<td>22</td>
<td>0.968</td>
<td>27</td>
</tr>
<tr>
<td>Fars</td>
<td>0.927</td>
<td>9</td>
<td>0.976</td>
<td>18</td>
<td>South kerman</td>
<td>0.877</td>
<td>23</td>
<td>0.976</td>
<td>14</td>
</tr>
<tr>
<td>Markazi</td>
<td>0.924</td>
<td>10</td>
<td>0.977</td>
<td>12</td>
<td>Semnan</td>
<td>0.876</td>
<td>24</td>
<td>0.975</td>
<td>21</td>
</tr>
<tr>
<td>Kohkilo</td>
<td>0.923</td>
<td>11</td>
<td>0.976</td>
<td>17</td>
<td>Sistan va Blochstan</td>
<td>0.875</td>
<td>25</td>
<td>0.976</td>
<td>15</td>
</tr>
<tr>
<td>Zanjian</td>
<td>0.922</td>
<td>12</td>
<td>0.978</td>
<td>11</td>
<td>Esfahan</td>
<td>0.867</td>
<td>26</td>
<td>0.971</td>
<td>25</td>
</tr>
<tr>
<td>Hamadan</td>
<td>0.918</td>
<td>13</td>
<td>0.980</td>
<td>3</td>
<td>Chaharmahal bakhtiyari</td>
<td>0.854</td>
<td>27</td>
<td>0.966</td>
<td>28</td>
</tr>
<tr>
<td>Ghazvin</td>
<td>0.918</td>
<td>14</td>
<td>0.974</td>
<td>22</td>
<td>Yazd</td>
<td>0.789</td>
<td>28</td>
<td>0.970</td>
<td>27</td>
</tr>
</tbody>
</table>

Source: Agricultural Ministry of Iran and research calculations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Describe</th>
<th>Number of observation</th>
<th>Mean</th>
<th>S.D</th>
<th>min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUP</td>
<td>Cost efficiency (Parametric method)</td>
<td>28</td>
<td>0.908</td>
<td>0.0412</td>
<td>0.789</td>
<td>0.995</td>
</tr>
<tr>
<td>EUNP</td>
<td>Cost efficiency (LOWESS method)</td>
<td>28</td>
<td>0.976</td>
<td>0.004</td>
<td>0.966</td>
<td>0.986</td>
</tr>
</tbody>
</table>
CONCLUSION

Our purpose in this study was survey and comparative parametric and partially non-parametric methods in wheat production cost efficiency evaluation in Iran. We used specification and estimation stochastic frontier cost function with panel data. With attention to quality of data in the agricultural sector of Iran, the results of efficiency measurement by specification and estimation cost function is more suitable than by specification and estimation production function. Because input prices instead of inputs were used and quality of price data is higher than quality of input data at province levels in cost function estimation. The results show that agricultural pesticides price hasn’t significant effect on frontier cost production. Share of agricultural pesticides is nugatory. Agricultural pesticides used when the farm is infected by disease therefore pesticides consume and yield per hectare are relevant and these reasons can be accounting for insignificant agricultural poison price.

In these two methods, wheat farmers Khuzestan province have the highest cost efficiency. Khuzestan province has vast fields as well as plenty of water for agriculture. The farm size in this province is high. Economics of scale seems to be one of the reasons for highest cost efficiency in Khuzestan province. Detailed studies in this field are offered and the main reasons for the high cost efficiency of irrigated wheat production in the province are found.

Parametric and nonparametric methods for estimating the model results indicate that from the point of view of significant coefficients and differentiated production structure, parametric approach in each province is more appropriate. Therefore, in similar studies, especially when panel data is used one should be more discreet to use partially non-parametric models.

Parametric and nonparametric methods show that cost efficiency of wheat production in Iran is suitable and over 90%. And this result emphasized on the fact that if farmers were given the facilities they would effectively use it.

It is necessary that department of agricultural promotion and education organize their forces and facilities based on the ranking in terms of cost efficiency of every province. So that in whatever province where the efficiency is low, agricultural education and knowledge is needed to improve further.

REFERENCES