Efficiency Series Paper 05/2004

Differences in Productivity Between Efficient and Inefficient Firms

Antonio Álvarez y Luis Orea

Departamento de Economía

Universidad de Oviedo

Available online at: www.uniovi.es/economia/edp.htm
Differences in Productivity between Efficient and Inefficient Firms

Antonio Álvarez y Luis Orea ♠

Efficiency Series Paper 05/2004

Abstract: Standard microeconomic theory does not deal with the behavior of inefficient firms. However it would be useful to know what happens to an inefficient firm if this firm grows or the effects of policy changes. In this paper we develop a model that allows the different behavior of efficient and inefficient firms to be compared.

Keywords: Productivity, inefficiency, cost function.
1. Introduction

Standard economic theory deals with the behavior of inefficient firms. Microeconomic models assume that the parameters of the production (cost, profit,...) frontier are shared by the inefficient firms. This assumption has an important drawback: it gives the same predicted value for firms regardless of their efficiency level. However, if a firm is inefficient at a particular period of time, then its behavior has been different from the behavior of the efficient firm. Thus, it is not reasonable to expect that if both firms increase size by the same amount then both would experience the same change in the dependent variable. It would be of interest, therefore, to know what happens to an inefficient firm when this firm grows or when an economic policy is implemented.

Despite the lack of theoretical attention that has been given to the behavior of inefficient firms, since Farrell (1957) a lot of effort has been devoted to the empirical identification of inefficient firms. This vast literature has mainly been concerned with measuring inefficiency, and very few papers have tackled the theoretical underpinnings of inefficiency analysis\(^1\). One strand of this literature is devoted to the explanation of inefficiency\(^2\). In these models inefficiency is considered as a function of exogenous variables, thus providing the basic framework to treat inefficient behavior separately.

The objective of this paper is to measure productivity in the context of a model that allows efficient and inefficient firms to behave differently. For this purpose we use a model recently developed by Orea and Kumbhakar (2004) which allows us to calculate total factor productivity (TFP) growth separately for efficient and inefficient firms. The capabilities of the model are illustrated with an empirical application that uses panel data from a sample of Spanish dairy farms.

The model has important implications for empirical research. The analysis of inefficient behavior is of great importance in developing economics insofar as the concept of inefficiency is at the core of many development issues. One of these issues is the well-known relationship between efficiency and farm size. Schultz’s (1964) argument that small farms are ‘poor but efficient’ was the origin of a large body of literature. However, some of the empirical studies that tried to test the validity of this hypothesis have been

\(^1\) A notable exception is Bogetoft (2000) and the references therein.
based on proxies for efficiency measures, such as productivity per acre (as in the influential paper by Sen (1962)). One advantage of our model is that it easily allows testing whether inefficient farms move towards or away from the frontier when they grow, i.e., whether they become more or less efficient.

The paper is organized as follows. Section 2 deals with the modeling of inefficiency. Section 3 develops TFP in the presence of inefficiency. Section 4 describes the data and the empirical model. Section 5 reports the estimation and results. Section 6 contains a summary and some conclusions.

2. Modeling inefficiency

The starting point of the model is a stochastic cost frontier (Aigner, Lovell and Schmidt, 1977). The model can be written as:

$$C_{it} = C^*(y_{it}, w_{it}, x_{it}, t, \beta) \exp(v_{it} + u_{it})$$  \hspace{1cm} (1)$$

where $C_{it}$ is the observed cost, $C^*(\cdot)$ is minimum cost, $y_{it}$ is the output level, $w_{it}$ is a vector of variable input prices, $x_{it}$ is a vector of quasi-fixed inputs, $t$ is a time trend, $v_{it}$ is a random disturbance term which is assumed to have zero mean and constant variance $\sigma_v^2$, and $u_{it}$, which accounts for cost inefficiency, is a non-negative random disturbance (in order to be on or above the cost frontier).

The estimated coefficients in (1) can be used to predict the behavior of each efficient firm as follows:

$$\hat{C}_{it} = C^*(y_{it}, w_{it}, t, \hat{\beta})$$  \hspace{1cm} (2)$$

However, this equation is only able to predict behavior on the frontier. That is, it ignores the possibility that inefficient firms behave differently. This point is illustrated in Figure 1, where there are two firms which produce the same output $y_0$ but firm A is efficient since it is located on the average cost frontier, while firm B is cost inefficient.

---

2 This literature is summarized in Kumbhakar and Lovell (2000).
The question is what will happen to these two firms when both increase output by the same amount.

![Figure 1. Behavior of efficient and inefficient firms](image)

If both firms increase output to \( y_1 \), model (2) predicts that firm A will move to A' but provides no information as to where will firm B go. The reason is that in model (1) the inefficiency term (u) takes a positive value but otherwise carries no information about the behavior of inefficient firms. In order to be able to predict B's movement more information is needed. This information can be obtained if some structure is imposed on the inefficiency part of the model so that the movement of B may be identified. The simplest case is to assume that inefficiency is time invariant. In this case B would move to B', where the inefficiency level is the same as in the initial situation. Therefore, the missing piece of information is the variation of inefficiency.

The first papers that attempted to model the inefficiency term in a stochastic frontier function were Kumbhakar, Ghosh and McGuckin (1991), Reifschneider and Stevenson (1991) and Huang and Liu (1994). These models allow inefficiency (u) to vary systematically with some exogenous variables (z). Their approach consists of making the mean of the distribution of the inefficiency term depend on a set of exogenous variables. These models were originally developed for cross-sectional data. Battese and Coelli (1995) extended this approach to accommodate panel data. As in the previous models, the inefficiency term is assumed to follow a truncated normal
distribution where the mean of the pre-truncated distribution depends on some exogenous variables. That is,

\[ u_{it} \sim N^+ (\mu_{it}, \sigma_u^2) , \quad u_{it} \geq 0 , \quad \mu_{it} = \delta \cdot z_{it} \]  

(3)

where \( z_{it} \) is a vector of explanatory variables. The coefficient vector associated with \( z_{it} \) measures changes in the mean of the pre-truncated distribution due to changes in \( z_{it} \). To predict the effect of \( z_{it} \) on the dependent variable we can analyze the marginal effect of \( z_{it} \) on the unconditional expectation of inefficiency, that is, \( \partial E(u_{it})/\partial z \).

An important feature of the Battese and Coelli (1995) model, which is shared by other models where inefficiency depends on exogenous variables, is that it assumes independence over time of the efficiency term. That is, a firm observed in two periods is treated as two different firms. This assumption does not allow us to estimate the inefficiency level consistently since its variance does not vanish as the sample size increases\(^3\). The alternative models, on the other hand, yield quite cumbersome expressions for the marginal effect of \( z_{it} \) on inefficiency.

For these reasons, we propose using the model introduced by Orea and Kumbhakar (2004), which overcomes these problems. In particular, we propose to directly model the inefficiency term by specifying \( u_{it} \) as the product of a function of some exogenous variables \( z \) and a non-negative, time-invariant but firm-specific inefficiency term, \( u_i \). That is, the proposed model is the following:

\[ u_{it} = g(z_{it}, \delta) \cdot u_i = \exp(\delta^\prime z_{it}) \cdot u_i , \quad u_i \geq 0 \]  

(4)

where \( \delta = (\delta_1, \ldots, \delta_k) \) are parameters and \( z_{it} = (z_{it1}, \ldots, z_{ikt}) \) is a vector of \( k \) variables that are assumed to affect inefficiency (e.g. size, time, form of ownership, managerial characteristics, etc.).

This specification yields several other parametric functions proposed in the literature as special cases. If \( T=1 \), this model collapses to the cross-sectional model introduced by Simar, Lovell and Van den Eeckaut (1994). If \( z \) is a time trend or a vector of time

\(^3\) Detailed discussions of this issue can be found in Schmidt and Sickles (1984) and Greene (1993).
dummy variables, it transforms into a homoskedastic panel data model with time-varying efficiency. In particular, if $\delta$ is a scalar and $z_{it} = (T-t)$ we get the specification proposed by Battese and Coelli (1992). If $\delta$ is a $1x2$ vector and $z_{it} = (t, t^2)$ we get a specification similar to that proposed by Kumbhakar (1990). Finally, if $\delta$ is a $1xT$ vector and $z_{it}$ is a set of $T$ time-dummy variables, we get the specification proposed by Lee and Schmidt (1993).

Some appealing features of the model in (4) are worth highlighting. First, since the inefficiency term is developed in a panel data framework, inefficiency is correlated over time, i.e. $\text{cov}(u_{it}, u_{it-1}) \neq 0$. This feature allows us to estimate the inefficiency level consistently when $T \to \infty$. In addition, since the inefficiency term is modeled as the product of a deterministic function of exogenous variables and a stochastic term, the marginal effect on inefficiency of a change in $z$ is separable into a deterministic and a stochastic component. In particular, the marginal effect on inefficiency of a change in $z$ is given by the following derivative:

$$\frac{\partial E(u_{it})}{\partial z_{it}} = \frac{\partial g(z_{it}, \delta)}{\partial z_{it}} E(u_{i})$$

(5)

3. Total Factor Productivity of inefficient firms

In order to analyze differences between efficient and inefficient firms we use the concept of Total Factor Productivity (TFP), which is a common measure of firm performance. TFP growth can be measured as (minus) the growth in average cost after controlling for the growth in input prices. From equation (1), total factor productivity growth for an efficient firm is given by (see, for instance, Denny, Fuss and Waverman, 1981):

$$\dot{\text{TFP}} = -\frac{\partial \ln C^*}{\partial t} + (1 - \varepsilon_y) \dot{y}$$

(6)

where a dot over a variable indicates a growth rate, $\varepsilon_y$ is the elasticity of minimum cost with respect to output, and the subscripts denoting firm $i$ in period $t$ have been dropped for notational ease. The first term measures shifts in the cost function, which are conventionally attributed to technical change. The second term measures movements
along the cost function when output expands over time and the technology exhibits increasing or decreasing returns to scale\(^4\).

Equation (6) shows how TFP change can be calculated from the estimated cost frontier parameters. However, since this measure is obtained after making some straightforward manipulations on the cost frontier, it does not include any inefficiency effect. Therefore, equation (6) predicts the same productivity performance for both efficient and inefficient firms. The productivity of an inefficient firm might be, however, quite different from that of an efficient farm when inefficiency changes over time. To allow for these differences, we write productivity growth for an inefficient firm as:

\[
\frac{d\text{TFP}_u}{dt} = \text{TFP} - \frac{dE(u)}{dt}
\]  

(7)

where the second term on the right hand side of the equation measures changes over time in expected cost inefficiency\(^5\).

In equation (1) the marginal effect \(dE(u)/dt\) is assumed to be zero, since no information about the behavior of inefficiency is incorporated. Therefore, using a standard stochastic frontier model precludes the study of differences in productivity among efficient and inefficient firms. Therefore, we need to relax the restriction \(dE(u)/dt=0\). Incorporating exogenous variables as determinants of firm’s inefficiency allows expected inefficiency to vary systematically with some exogenous variables. Thus, given the specification of \(u\) in equation (4), the marginal effect of time on the expected inefficiency is:

\[
\frac{dE(u_i)}{dt} = \sum_{k=1}^{K} \frac{\partial g(z_i, \delta)}{\partial z_k} \cdot E(u_i) \cdot \dot{z}_k
\]  

(8)

\(^4\) The second term depends on the degree of local returns to scale measured as one minus cost elasticity with respect to output. Increasing scale economies are indicated by a positive value and decreasing economies by a negative value. Hence, an output expansion leads to an increase (decrease) in total factor productivity when increasing (decreasing) returns to scale exist.

\(^5\) This specification of productivity growth for an inefficient firm is quite similar to the one introduced by Bauer (1990), where the efficiency term is defined in levels instead of expected values.
This derivative indicates that changes in the inefficiency term rely on the magnitude of the change ($\Delta z$), the “marginal” effect of $z$ on the deterministic part of the inefficiency term ($\Delta g/\Delta z$), and the initial inefficiency level ($u$).

4. Data and empirical model

This study uses data from a group of 96 dairy farms located in Northern Spain that are enrolled in a voluntary Dairy Cattle Management Program. We have data on these farms for a period of four years, 1999-2002.

We estimate a translog cost frontier, where $u_{it}$ is specified as in (4). The dependent variable is operating costs. The output (Milk) is measured as liters of milk produced. As for the input prices, we include the average annual price that each producer pays for foodstuffs (Feed). Land, measured in hectares, is included as a quasi-fixed input. The empirical model can thus be written as follows:

$$
\ln C_{it} = \beta_0 + \beta_1 D_t + \beta_y \ln(\text{milk})_{it} + \beta_p \ln(\text{Pfeed})_{it} + \beta_L \ln(\text{land})_{it} + \beta_{yy} 1/2 \cdot \ln(\text{milk})_{it}^2 + \beta_{pp} 1/2 \cdot \ln(\text{Pfeed})_{it}^2 + \beta_{LL} 1/2 \cdot \ln(\text{land})_{it}^2 + \beta_{yp} \ln(\text{milk})_{it} \ln(\text{Pfeed})_{it} + \beta_{yl} \ln(\text{milk})_{it} \ln(\text{land})_{it} + \beta_{pl} \ln(\text{Pfeed})_{it} \ln(\text{land})_{it} + v_{it} + \exp[\delta_y \ln(\text{milk})_{it} + \delta_{yy} 1/2 \cdot \ln(\text{milk})_{it}^2 + \delta_p \ln(\text{Pfeed})_{it}^2 + \delta_{pp} 1/2 \cdot \ln(\text{Pfeed})_{it}^2 + \delta_{yp} \ln(\text{milk})_{it} \ln(\text{Pfeed})_{it} + \delta_{yl} \ln(\text{milk})_{it} \ln(\text{land})_{it} + \delta_{pl} \ln(\text{Pfeed})_{it} \ln(\text{land})_{it} + \delta_{l} t] u_{it}
$$

where the time dummy variables $D_t$ capture not only neutral technical change but also the effect of other input prices which are common to all producers, and where the vector of inefficiency determinants includes milk, which is taken as a measure of farm's size, the feed price, and a time trend.

---

6 We have not included other quasi-fixed inputs due to multicollinearity problems.
5. Estimation and results

The cost function in (9) can be estimated by maximum likelihood assuming that \( u_i \) is a non-negative random variable that follows a one-sided distribution. In our empirical application we have assumed a half-normal distribution for the inefficiency term. The model was estimated using GAUSS and the results can be seen in Table 1. Since the explanatory variables in the cost function have been normalized by their geometric means, the first order coefficients can be interpreted as cost elasticities evaluated at the sample geometric mean. The estimated variable cost function satisfies the regularity conditions at this point. The estimated coefficients of the variable input price and of the quasi-fixed input have the expected sign and are statistically significant. The time effects are negative and grow in absolute value over time, which can be interpreted as evidence of increasing technical progress.

Turning to the inefficiency term, the estimates show that inefficiency tends to decrease with feed price, which is a reasonable result if higher prices reflect better feed quality. The coefficient on the time trend is positive and statistically significant indicating that, holding output and feed price constant, inefficiency has increased over time.

Note that, applying Shephard’s Lemma, the elasticity of cost with respect to feed price is the cost share of feed. If inefficiency decreases when the feed price rises, this means that an inefficient farm uses proportionately less feed (but uses more of other inputs) than an efficient farm\(^7\). As in a process of natural selection, it is reasonable to assume that only the efficient firms will survive in the face of adverse conditions. Thus, the disappearance of a significant proportion of the firms comprising a given sector (due to, for instance, a process of modernization and liberalization) can induce changes both in the form of a reduction in the number of firms as well as in the intensity with which the sector uses the different inputs, which will have environmental implications given that livestock use is contaminating (e.g. Innes, 2000).

\[ S_k = S_k^* + \frac{\partial u}{\partial \ln w_k} \]

\(^7\) Note that inefficiency \( u \) depends on an input price, \( w_k \), so differentiating equation (9) we get
Table 1. Parameter estimates of the cost frontier

<table>
<thead>
<tr>
<th>Frontier</th>
<th>Coefficient</th>
<th>Std. Dev.</th>
<th>Inefficiency Term</th>
<th>Coefficient</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>10.4258</td>
<td>0.0259</td>
<td>Ln(milk)</td>
<td>-0.6918</td>
<td>0.214</td>
</tr>
<tr>
<td>Ln(milk)</td>
<td>1.2587</td>
<td>0.0507</td>
<td>Ln(milk)^2</td>
<td>0.0592</td>
<td>0.3934</td>
</tr>
<tr>
<td>Ln(Pfeed)</td>
<td>0.7508</td>
<td>0.1395</td>
<td>Ln(Pfeed)</td>
<td>-2.0231</td>
<td>0.7944</td>
</tr>
<tr>
<td>Ln(land)</td>
<td>-0.1661</td>
<td>0.0403</td>
<td>Ln(Pfeed)^2</td>
<td>-8.9795</td>
<td>5.6889</td>
</tr>
<tr>
<td>Ln(milk)^2</td>
<td>-0.4451</td>
<td>0.1231</td>
<td>Ln(milk) Ln(Pfeed)</td>
<td>-3.8516</td>
<td>1.647</td>
</tr>
<tr>
<td>Ln(Pfeed)^2</td>
<td>2.3203</td>
<td>0.8176</td>
<td>T</td>
<td>0.0983</td>
<td>0.0506</td>
</tr>
<tr>
<td>Ln(land)^2</td>
<td>-0.3713</td>
<td>0.1207</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(milk) Ln(Pfeed)</td>
<td>0.9988</td>
<td>0.388</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(milk) Ln(land)</td>
<td>0.6162</td>
<td>0.0718</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(Pfeed) Ln(land)</td>
<td>-0.2079</td>
<td>0.2511</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2000</td>
<td>-0.0438</td>
<td>0.0187</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2001</td>
<td>-0.0743</td>
<td>0.0245</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2002</td>
<td>-0.1224</td>
<td>0.0317</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^2 = \sigma^2_v + \sigma^2_u$</td>
<td>0.0487</td>
<td>0.0104</td>
<td>Observations</td>
<td>384</td>
<td></td>
</tr>
<tr>
<td>$\lambda = \sigma_u / \sigma_v$</td>
<td>1.680</td>
<td>0.111</td>
<td>Log-Likelihood</td>
<td>142.223</td>
<td></td>
</tr>
</tbody>
</table>

What is the role of farm size implied by this model? The answer to this question depends on the part of the cost function we are considering. In the frontier part the cost elasticity with respect to output, evaluated at the sample mean, is greater than one, indicating the existence of decreasing returns to scale on the frontier. Therefore, the average cost of an efficient farm increases with farm output, holding land fixed. On the other hand, the coefficient associated with milk output in the inefficiency term is negative and significantly different from zero, implying that inefficiency decreases when farm production rises.

This result is quite important if we want to measure the consequences on farms' competitiveness of policy measures which, directly or indirectly, encourage farm growth. One example of this kind of policy is a voluntary abandonment scheme that increases farm's size as the quota freed up is allocated among the remaining farms. The effect of this measure on average costs depends on whether the inefficiency effect is large enough to compensate the decreasing returns of the technology. In that
case, the average cost of an inefficient firm may actually decrease, increasing farms’ competitiveness as a whole.

From the estimated parameters we have calculated the size elasticity measured as the ratio of marginal cost to average cost (Hanoch, 1975). This can be done using either the observed cost or, assuming that each farm is efficient, the efficient cost\(^8\). Thus, for each observation we get two size elasticities. Figure 2 shows that, for a given size, the size elasticity for inefficient farms is smaller than for efficient ones. In fact, the size elasticity on the frontier is, on average, 10% higher than the observed elasticity. It is noteworthy that the cost-reaction of inefficient farms when they increase their output levels is, in general, quite different (in some cases more than 50%) from an efficient farm that increased its output level in the same proportion. In order to properly forecast the consequences of a particular policy measure, these different reactions should therefore be taken into account.

We now use the parameter estimates in Table 1 to calculate TFP growth for both efficient and inefficient farms, using equations (6), (7), and (8). The results are presented in Table 2. Using the numbers in Table 2, we have constructed a cumulative index for the TFP of efficient farms (labeled TFP growth), which ignores changes in efficiency. This cumulative index as well as its decomposition into technical change and a scale effect are shown in Figure 3. This figure shows that productivity increased only moderately, even though there was considerable technical progress. This is explained by a negative scale effect (there are diseconomies of scale and farms have increased production) which has partially offset the positive effect of technical change. On the other hand, the productivity of inefficient farms, (labeled TFP\(_U\) growth and also shown in Figure 3) has increased more than in the case of efficient farms since their inefficiency has decreased over the period (dE(u)/dt is negative).

\(^8\) While the efficient size elasticity is obtained by differentiating the cost frontier, the observed size elasticity is measured as follows:

\[
\varepsilon_y = \frac{\partial \ln C}{\partial \ln y} = \frac{\partial \ln C^*}{\partial \ln y} + \frac{\partial u}{\partial \ln y}
\]
Table 2. Decomposition of TFP growth

<table>
<thead>
<tr>
<th></th>
<th>Frontier part</th>
<th>Inefficiency part</th>
<th>dE(u)/dt</th>
<th>TFP growth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Technical Change</td>
<td>Scale Change</td>
<td>TFP growth</td>
<td>Milk</td>
</tr>
<tr>
<td>2000/1999</td>
<td>4.38</td>
<td>-3.61</td>
<td>0.77</td>
<td>-2.45</td>
</tr>
<tr>
<td>2001/2000</td>
<td>3.05</td>
<td>-2.89</td>
<td>0.16</td>
<td>-1.32</td>
</tr>
<tr>
<td>2002/2001</td>
<td>4.81</td>
<td>-3.31</td>
<td>1.50</td>
<td>-1.88</td>
</tr>
</tbody>
</table>
The evolution of efficiency and its determinants over time is shown in Figure 4 using cumulative indexes.
This figure allows us to establish that the improvement in efficiency can be mainly explained by increases in farms’ output and feed prices, which have compensated the reduction in efficiency over time due to the effect of other variables as captured by the time trend. In summary, our results clearly show that the incorporation of efficiency changes can affect the evaluation of farms’ productivity.

6. Conclusions

In this paper we use a model that allows for differences in the behavior of efficient and inefficient firms. This model differs from previous ones in that it fully exploits the panel data nature of the data. We believe that this model can be very useful for policy purposes since the standard microeconomic models are only able to predict the behavior of efficient firms, which usually are a small fraction of the sample.

Some of the results of the present paper have relevant policy implications. In particular, we show that size has two different effects on farm productivity. On the one hand, since there are diseconomies of scale in the technology, an increase in size reduces productivity. On the other, inefficiency decreases with size and farms’ competitiveness (measured in terms of average costs) may therefore increase as a whole if a policy measure that increases size, and in turn efficiency, is implemented. Therefore, a detailed knowledge of the consequences of measures that may change efficiency (e.g. farmer training) or size (e.g. voluntary abandonment schemes) is important for policy design in this sector.
References


