**A Multilevel Approach to Measuring** **Efficiency**

# Abstract

The conventional single level stochastic frontier model ignores the structure of the data resulting in the falsification of efficiency scores. Accordingly, the multilevel models, a special case of random coefficients models, take into account, the structure of data but are not frontier. These models are extremely uncommon to researchers, more so, their applicability. Researchers in both theory and practice have attempted to incorporate these models in their respective fields but with little success. In a bid to show the effects of ignoring the structure of data in efficiency analysis, clustered data is generated, partly in Rstudio and in partly in Limdep. A multilevel stochastic frontier model is proposed. The proposed model is run in Limdep using simulated dataset. The Three versions of the conventional single level stochastic frontier model are also run in the same software and with the same simulated dataset. Results of the proposed model are compared to those of the conventional models. The comparisons carried out show that ignoring the structure of data in efficiency analysis distorts the efficiency levels as well as the rankings of the subjects in the analysis. Efficiency levels from the developed model are significantly higher than those of any of the three versions of the conventional model, suggesting that efficiency levels being estimated in hierarchical organizations like municipalities are underestimates of reality. The adoption of the proposed model will in these situations will rectify this anomaly.

**keywords:** frontier, simulate, random coefficients, copula, clustered data.

# Introduction

(Aigner et al: 1977) and (Meeusen and van den Broeck; 1977) independently proposed two-component error term stochastic frontier model, which was followed by various attempts to evaluate the efficiency of production processes and cost functions. In this context, a Stochastic Frontier is defined as the function *f*(**x***i,β*) + **v***i*. Given a bundle of inputs **x***i*, the actual production level is *yi* ≤ *f*(**x***i,β*) + **v***i*, (Aigner et al; 1977). For a cost function, therefore, actual cost is *yi* ≥ *f*(**x***i,β*) + **v***i*, where *f*(**x***i,β*) is the mean production level and **v***i* is the symmetric disturbance term.

Developments in this model include the use of panel data so that it becomes possible to calculate individual firm efficiency as well as separating time-invariant and persistent inefficiency. Attempts have been made to separate individual heterogeneity from inefficiency. Some heterogeneity is due to clustering. The conventional Stochastic Frontier Model neglects the hierarchical structure of data, (Siciliani; 2006). Researchers such as (Johns, 2006), (Lordan, 2009) and (Martha, 2011) have studied efficiency estimation in cases where the units of analysis are clustered. Each of them attempted to use multilevel models in efficiency estimation.

 This paper examines the extent to which multilevel models have been used in efficiency analysis. A multilevel stochastic frontier is presented and simulated data is utilized to compare inefficiency levels from the proposed model to those from three other stochastic frontier models. These three are: (1) a single-level frontier that ignores information from other levels of the data structure and has fixed coefficients; (2) a model that also ignores information from other levels of the data structure but has random coefficients; and (3) a third model that accounts for the structure of the data through aggregation.

A study to explore the factors that affect the production of rice in the state of Uttar Pradesh, India was carried out by (Chandel et al, 2022). Realizing that the state of Uttar Pradesh is divided into four zones, the researchers used the appropriate multi-stage sampling technique to come up with their sample of 3200 farms. From each of the four zones, four towns were randomly chosen, and from each town, 200 farms were selected. The independent variables used in the study included crop type, seed quality, seed variety, water availability, environmental aspects, labor, risk management, soil fertility, farmer’s educational level, farmer’s experience, marketing information, fertilizer quantity, quality of pesticides and tenure status.

 Not all these variables belong to the same level. For example, soil fertility is a zone level (level 3) variable while educational level and experience are at farmer level, (level 1). Environmental aspects could be at town (level 2) or at zonal, (level 3). The researchers acknowledge that there were differences in production levels among the four zones and also among the farmers. However, they went on to feed the variables into the standard multiple linear regression model which takes no cognizance of the nesting structure of the data. Regardless of recognizing the hierarchical nature of the data through multi-stage sampling, the researchers had no alternative but to use conventional single-level stochastic frontier model.

The results of this research revealed that labor, education, irrigation, and seed-type had a significant positive effect on productivity. Surprisingly, experience and tenure had a negative impact to yield, though the effect was insignificant. The mean technical efficiency level was 72%. There are chances that this mean could have been higher had the data been appropriately modelled.

The goal of this research is to develop an appropriate model for hierarchical data. The remainder part of the paper is structured as follows: Section 2 looks at the application of multilevel models in efficiency estimation. The multilevel stochastic frontier model is proposed and discussed in section 3. In section 4, data is simulated and used to estimate efficiency using the four models mentioned above. Section 5 discusses simulation results and section 6 presents the conclusion.

# Multilevel Models and their Application in Efficiency Estimation

(Gelman, 2007) defines a multilevel model as a regression model in which the regression coefficients are defined by a probability model. According to this definition, varying coefficients and the definition of such coefficients in the form of a model, are the key features of these models. The models are known by many other names in different areas. Some of these names, together with their respective areas of use are: (1) Hierarchical models, Statistics, (Harville, 1977), (2) random coefficients models, (3) Econometrics, (Swamy, 1970), random-effects models, (Laird and Ware, 1982) and (4) mixed effects models, Biostatistics (Goldstein, 1986).

These models are an extension of regression in which data are structured in groups and the coefficients of the models are allowed to vary by groups, (Paterson and Goldstein, 1991, Twisk, 2006). The feature that makes them special is that their coefficients vary by groups or clusters. A unique characteristic of multilevel models is that there is at least one equation at each of the levels of the structured data. These models take many forms. Some are random intercept, others are random slope, with other being random intercept random slope.

Liouaeddine and Bijou (2021) used the model in its random intercept form to analyze the performance of students from Morocco in Mathematics. The independent variables in this research included gender, age and personal computer ownership, at student level and school location, parents’ educational level and socio-economic status at school level.

The model used in the analysis is

  *(1)*

Where i is the student and j is the school.

A rearrangement of the terms of the equation shows that the equation is a multilevel one with only the intercept varying. The implication of this model is that the school-level variables only affect the intercept. Students from different schools start at different levels in Mathematics, but they develop at the same rate. All the variables in the model were found to be significant. Boys were proved to do better in Mathematics than girls. Students with higher socio-economic status proved to be better in Mathematics than those with low economic status. The analysis of the data also showed that the educational level of parents has a positive impact on the performance of students in Mathematics.

To determine the best approach in the analysis of statistical data, it is necessary to consider the structure of the data. Some data is clustered. Clustered data is data that is classified into distinct groups commonly referred to as clusters, (Galbraith et al., 2010) These data are distinguished from single-level data sets by the nesting of individual observations within higher level groups, (Heck et al, 2014). The effects of clustering, if ignored, result in the underestimation of the standard error, as more information is assumed to be provided by the data than what the data actually provides, (Thommai, 2019). Mundfrom and Schultz (2002) are among the researchers who carried out simulation studies to try to establish the effects of clustering on data and data analysis methods. They pointed out that practitioners do not employ hierarchical linear models mainly because they are not aware of them. They simulated data using several intra-correlation coefficients (ICC). The data was analyzed using both multiple linear regression (MLR) and hierarchical linear models (HLM). Their discovery was that there is no problem with regression coefficients produced by these two techniques as they are always equal regardless of the ICC. The two researchers mentioned above agree with others in the area that standard errors are a cause for concern. The single-level model, which does not account for the structure of the data, tends to underestimates the standard error. McNeish (2014) notes that standard errors are a function of the sample size and as mentioned by Thommai (2019), clustered data provides less information than what the sample size assumes. For such data several observations in a cluster give the same information. If, for example, all of the observations in a cluster are identical, no more information is given by these observation than given by only one observation. As a result the effective sample size for that cluster is one and not the cluster size.

Since hierarchical models account for clustering effects while multiple linear regression does not, the sample sizes used by the two techniques differ. Multiple linear regression maintains the original sample size, while hierarchical linear models make use of the effective sample size, which may be much less than the original sample size, depending on the ICC. As functions of the sample size, the standard errors of the two techniques are made to differ.

McNeish (2014) generated data using a two-level random intercept model. The data had only one level 1 variable. The slope of this level 1 variable was fixed and there were no level 2 variables hence no interaction effects. Like Mundfrom and Schultz (2002), McNeish (2014) manipulated the ICC. Two models, MLR and HLM were run. Results obtained were similar to those obtained by Mundfrom and Schultz (2002) and many others who carried out similar studies.

A simulation study by Thommai (2019) examined clustering effects on model parameter estimates precision. In a clustered design the members of the sample are put into two or more distinct groups, with each member belonging to only one group. This causes the data to have two or more levels, referred to as individual and group, (Galbraith et al., 2010). This hierarchical or multilevel feature of the data causes members within the same group to be more similar than those between groups, which implies non-independence of within group members, (Galbraith et al., 2010). In the analysis of such data, the intra correlation coefficient, which is the quantitative measure of variance that is accounted for by the clustering effects, plays an important role. Noted in this study is the fact that by ignoring clustering effects, the data is assumed to contain more information than it actually does. This results in overestimation of the precision and false statistical significance conclusions.

An assertion was made that the intra-correlation coefficient can take a value that is either positive or negative. Thommai (2019) derived the causes and interpretation of negative correlation from Pryseley et al. (2011) who described such correlations as resulting from dissimilarities within groups. They, Pryseley et al. (2011), also noted that negative correlation can be a result of unstable variance or covariance. Other factors that have been found to promote negative intra correlation coefficient are the sample size and the cluster size. These must not be small, (Pryseley et al., 2011). An observation by Thommai (2019) was that the models, ANCOVA, LME and LM falsely reflect no clustering whenever there is a negative correlation. Thommai (2019) did not propose a new model but instead ran a fixed effects model several times varying the size of the correlation coefficient. In addition, the model used by Thommai (2019) was random coefficients, not multilevel. Neither was the model frontier. The model that we propose here is a multilevel stochastic frontier. The model id both frontier and has random coefficients that vary by level.

Galbraith et al. (2010), cited in Thommai (2019) point out that clustered data pose challenges to researchers as methods that handle such data are uncommon and not widely understood. As such, researchers, when the use such data, either ignore the structure of the data and treat all observations as independent, or have resort to aggregation (Galbraith et al., 2010). Disaggregation is also common to researchers. Nevertheless, this choice of analyzing data at one level over another has resulted in fallacies.

In an attempt to separate individual heterogeneity from inefficiency (Greene, 2005) proposed what he called the “True fixed effects” model (below):

 *(2)*

The term is introduced for purposes of absorbing all heterogeneity assumed to be mistaken for inefficiency in the estimation process. This model is similar to the random intercept one. For the random coefficients model, the individual *αi*’s are not calculated. It is only their average which is calculated. For the above model, however, all the *αi*’s are calculated. This becomes a challenge when the number of study units is large.

Realizing the effect of clustering on results, several researchers in efficiency estimation have incorporated multilevel models in one way or the other. In a study aimed at establishing the extent to which the performance of Italian mutual-cooperative banks is affected by both geographical (provincial) and individual characteristics, Aiello and Bonanno (2015), recognized the tiered nature of the data and used multilevel modelling. Their dataset was panel so they proposed a three-level model. The three levels were time, bank and local market (province). The level 1 model **was** given by equation (3):

 (3)

where *ytij* is the estimated mutual-cooperative banks cost efficiency, *β*0*ij* are the intercepts, *β*1*ij* are the slope coefficients and *ϵtij* is the random error component. The time variable coefficient is *δij*. The time *t*, assumed values 2006 to 2011. The province is represented by *j* while *i* is for the individual bank.

The level 2 models and the level 3 models, from Aiello and Bonanno (2015), are given by equations (4) and (5) respectively:

 *(4)*

*and*

 *(5)*

The final model is obtained by substituting level 3 into level 2 and level 2 into level 1.

The variables at bank level included bank size, cost efficiency, cost income and equity. At provincial level, the variables included branches per municipality, market concentration on bank branches, market concentration on total assets and share of the top 3 banks. The empty multilevel model was run and it was found that 28.27% of bank heterogeneity i.e. efficiency, was explained by local markets. The individual banks themselves explained 28.11% and the remaining percentage was explained by the time effects.

The multilevel model was not used as a frontier in this case. It was only used to show the contribution of the factors at each of the three levels to the efficiency of the individual banks. The model had explanatory variables at level 1 only. The model proposed here is multilevel and frontier at the same time.

Multilevel models were used by Johnes (2006), in a study whose main theme was to determine whether or not the performance of universities depended on the technique used. The two techniques that were of interest are data envelopment analysis (DEA) and multilevel modelling (MLM). A dataset of graduates from universities in the United Kingdom was used in this research. The dataset had two levels, the student level at level 1 and the university level at level 2. The outcome variable of this research was the weighted degree class. The explanatory variables were age, gender, marital status and entry qualification. Like Aiello and Bonanno (2015), Johnes (2006) had explanatory variables only at the level of analysis. For the former, this was the bank level while for the latter, it was the student level. Johnes’ model is given in equation (6):

*yij* = *β*0 + *β*1*x*1*ij* + *β*2*x*2*tj* + *βkxkij* + *ϵij,*

*β*0*j* = *β*0 + *u*0*j* + *ϵij,*

 *β*1*j* = *β*1 + *u*1*j.* (6) (6)

Of interest in this part of the research were the university effects, which are the level 2 intercept errors, *u*0*j*. These errors were estimated using equation (7), (Johnes, 2006).

 *.* (7)

These effects were ranked and compared to those from DEA. Findings from this research showed that there was a very strong positive correlation between the rankings from the two techniques.

It is crucial to note that, although the researchers in theory and those in practice refer their models as multilevel or hierarchical, according to Bickel (2007), they are not. Bickel (2007) specifies what a multilevel model is. If there are no independent variables at the higher level to explain the variation in the random coefficient, then, the model is a general random coefficients model but not multilevel. All the models above have explanatory variables only at level 1, denying them the fit for multilevel. It is also surprising to note that Aiello and Bonanno (2015), who cite Bickel (2007), suggesting that they know the distinction between the general random coefficients model and the special case of it which is MLM, still mix the two.

Lordan (2009) developed a hierarchical frontier model that she used to estimate the efficiency levels of some group of health centers in Ireland. The two levels of the dataset were the centers at level 1 and the co-ops at level 2. This is because each center operates under a co-op and there are several centers under each co-op. Like Aiello and Bonanno (2015), panel data was used in this research. The data in this research had explanatory variables at both the center and co-op levels, making it hierarchical. Unfortunately, Lordan failed to create the implied interactions explained by Bickel (2007). Like in Bickel (2007) the researcher fails to reduce the model to the form it has here, a form that clearly shows the similarities between the general random coefficients and the special MLM. The model was kept in the state shown in equation (14) and it is not clear how the estimates of the parameters were obtained. It is the interaction term(s) that make MLM unique and these need to be clearly shown, as in equation (15).

Though the researcher compared the developed model to the standard stochastic frontier model, the comparison is not statistical in the sense that no hypothesis was tested to determine whether or not the differences between the efficiency levels of the various models considered were significant. In this study a model with interaction terms is proposed and hypothesis testing is carried out to compare the efficiency levels of the proposed model to those of currently existing models.

One of most recent publication on stochastic frontier analysis is Ngueni (2020). The researcher identified techniques that had not been applied to frontier modelling. These techniques include regular models, copula modeling and the nonparametric method of sieves. Ngueni (2020) discussed these models as new techniques in the frontier environment. Unfortunately, MLM was left out. The exclusion of MLM from the list of models that had not been used in efficiency estimation shows that these models (MLM) are in a worse situation in as far as application to frontier modelling is concerned.

# The Multilevel Stochastic Frontier Model

Equation (8) is a random coefficients model because the coefficients carry subscripts that suggest they vary. This means that we can build models that explain how these coefficients vary.

 (8)

If model (8) is accompanied by sub-models that describe how the coefficient vary, together they become a system as given by equation (9):

 (9)

The first of these 3 models is at level 1 whereas the other two are at level 2. A substitution of equation (9) into equation (8) produces equation (10) below:

 + (1*j*)(10)

The model that we propose here is frontier, implying that the error term *ϵij* is expressed as *vit* − *ui* so that the model is given by equation (11):

 + (1*j*)(11)

The system given by equation (9) is still not MLM because it does not have variables at level 2. When the level 2 equations take the form given in equation (13), the system then becomes multilevel, (Bickel, 2007). For simplicity, the model created here has only one explanatory variable at each of the two levels. Models (12) and (13) give the level 1 and level 2 equations with *x*1 and *x*2 as level 2 and level 1 variables respectively.

 + (1*j*)*.*

++ (12)

++ (13)

Substituting equation (13) into equation (12) results equation (14).

 + (1*j*)*.*  (14)

When like terms are grouped the resultant model reduces to equation (15)

 + (1*j*)*.*

# Methodology:

# Simulation

The data was generated using `Rstudio and Limdep. Two indicator variables, c and p, and four continuous variables, p1, x1, x2 and x12 were generated in Rstudio. The data was panel, with five observations coming from each of the three hundred individuals. This resulted in 1500 observations. Variable p1 was generated in such a way that it took three hundred values, with each value being repeated five times. Each set of five equal values was meant for each of the three hundred individuals described.

**Table 1: Part of the simulated data**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| N | Clus | Subj | p1 | x1 | x2 | x12 | y |
| 1 | 1 | 1 | 0.0064408 | 1.4611411 | 1.7929403 | 2.6197387 | 8.3723493 |
| 2 | 1 | 1 | 0.0064408 | 1.4611411 | 1.4132921 | 2.0650191 | 5.6674039 |
| 3 | 1 | 1 | 0.0064408 | 1.4611411 | 0.0336405 | 0.0491535 | 5.6331471 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 6 | 1 | 2 | 0.0078041 | 1.4611411 | 1.1396792 | 1.6652321 | 4.0780485 |
| 7 | 1 | 2 | 0.0078041 | 1.4611411 | 1.8505938 | 2.7039786 | 8.6422997 |
| 8 | 1 | 2 | 0.0078041 | 1.4611411 | 1.5282577 | 2.233 | 6.9672694 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 31 | 2 | 7 | 0.0069311 | 1.2048969 | 1.0227277 | 1.2322814 | 5.87605 |
| 32 | 2 | 7 | 0.0069311 | 1.2048969 | 1.3974195 | 1.6837464 | 4.0404052 |
| 33 | 2 | 7 | 0.0069311 | 1.2048969 | 1.8686962 | 2.2515862 | 7.1474498 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 36 | 2 | 8 | -0.003038 | 1.2048969 | 1.3494429 | 1.6259395 | 5.3051747 |
| 37 | 2 | 8 | -0.003038 | 1.2048969 | 2.4042436 | 2.8968656 | 10.3543182 |
| 38 | 2 | 8 | -0.003038 | 1.2048969 | 3.7420361 | 4.5087677 | 8.8004908 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 1498 | 50 | 300 | 0.0051074 | 1.58945 | -0.1470265 | -0.2336913 | -0.0538251 |
| 1499 | 50 | 300 | 0.0051074 | 1.58945 | 0.675273 | 1.0733127 | 4.2114045 |
| 1500 | 50 | 300 | 0.0051074 | 1.58945 | 1.2682358 | 2.0157974 | 69+-4.1118063 |

 The three hundred individuals were divided into fifty clusters of size six each. The variable x1 was simulated in such a way that there was one value for each of the fifty clusters. Each cluster had six individuals. With each individual having five observations, this equates to thirty observations per cluster. Variable x2 was normal, with 1500 observations. The above description means that p1 is a panel level variable, x2 is a level 1 variable and x1 is the level 2 variable. Two other variables created in Rstudio were *u*0 and *u*1. Each of these two was normal with mean zero and constant variance. These represented the level 2 errors in the model.

In addition to the variables created in Rstudio, three variables were generated in Limdep. These are u, v, and y. Variable v is has mean zero, constant variance and represents white noise. Variable u is half normal and represents inefficiency. The dependent variable, y, was generated as shown in equation (15). As already explained, equation (15) is obtained by substituting equations (13) into the random coefficients model (12). The resultant equation is then simplified so that it is clearly a random coefficients model but with a level 2 variable and an interaction term, as shown. Part of the generated data is shown in table 1. Variables *u*0*, u*1 and *v* are not shown here because of limited space. It is model (15) which we are proposing. It is a random coefficients model with cross-level interactions. Both the intercept and the slope of x2 (the level 1 variable) are random.

Before running the proposed model, however, three other models were also run in Limdep. These three were, the standard stochastic frontier model (single level), which ignores the structure of data and makes use of only x2, the level 1 variable, as the only explanatory variable, the random coefficients model with no explanatory variables at level 2, and one that recognized explanatory variables at level two but used disaggregation to model the data. The efficiency levels from all the four models were compared.

# Results and Discussion:

Data generation was performed several times, varying the mean and variance of x1 and x2. The results reported here were produced when the commands for *x*1 and *x*2 were

 and respectively. Similar results were obtained when the means and the variances of the two variables, x1 and x2 were changed.

Before the analysis, the data, *yijt* values, were generated in Limdep 11 using equation (15). *β*00 had value 3, and *β*01, *β*10 and *β*11 were each equal to 1. All other variables were generated as describe in the simulation section. The complete dataset was then used to generate the four frontier models, namely, model 1, a single-level model that ignores the structure of the data, model 2, which is a random coefficients model but has no explanatory variables at level 2, model 3, which uses aggregation to accommodate level 2 independent variables, and model 4, which makes use of the multilevel technique.

**Table 2: Descriptive Statistics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean | Std Dev | Min | Max |
| Model 1 (OLS) | 0.213943 | 0 .141309 | 0.000436 | 0.628775 |
| Model 2 (RC) | 0.235743 | 0.001209 | 0.001209 | 0.676815 |
| Model 3 (OLS) disaggregated | 0.282733 | 0.1591 | 0.004244 | 0.683397 |
| Model 4 (MLM) | 0.284654 | 0.168637 | 0.005884 | 0.702591 |

This study sought to establish whether or not the efficiency levels given by the proposed multilevel stochastic frontier are different from those produced when the structure of the data is ignored. Table 2 shows the descriptive statistics on the efficiency levels from the four models. From the table, model 4 has the largest maximum, largest minimum, hence largest mean efficiency levels. This suggests that ignoring the structure of the data, especially when there are explanatory variables at level 2, results in the underestimation of the efficiency levels.

**Table 3: Coefficients, Standard Errors and Probabilities.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Variable | Model 1 | Model 2 | Model 3 | Model 4 |
| Coefficients | Fixed | Constant | 4.60416 |  | 1.01054 |  |
| x1 |  |  | 2.49338 | 1.08704 |
| x2 | 2.39000 |  | 2.27909 |  |
| x12 |  |  |  | 0.94901 |
| Random | Constant |  | 4.62417 |  | 2.94146 |
| x2 |  | 2.32191 |  | 0 98919 |
|  |  |  |  |  |  |  |
| Std Error |  | Constant | 0.09832 | 0.12035 | 0.13307 | 0.22166 |
|  | x1 |  |  | 0.08163 | 0.13550 |
|  | x2 | 0.05344 | 0.04570 | 0.04186 | 0.11492 |
|  | x12 |  |  |  | 0. 07723 |
|  |  |  |  |  |  |  |
| P-Value |  | Constant | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
|  | x1 |  |  | 0.0000 | 0.0000 |
|  | x2 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
|  | x12 |  |  |  | 0.0000 |

Table 3 shows the estimates of coefficients, the standard errors and the p-values of the stochastic frontier model for each of the described model structures. Looking at the estimates of the coefficients of the variables in each of the four models, it is model 4 which has coefficients that agree with those used in the generation of the data. The implication here is that when building models to mimic data, it is important to take into account the structure of the data.

Each of the four models has a constant term and a coefficient of variable x2. Considering the standard errors of these two, it is clear from table 2 that model 4 has the largest, as supported by literature. The literature, as reviewed in points out that ignoring the structure of the data results in the underestimation of the standard error.

Although all the p-values here suggest that all coefficients of the four models are significantly different from zero, those of the other three models, that is, models 1 to 3, are significantly different from those that were used to generate the data. If we consider the constant term of model 3 which is 1.01054, we see that this is clearly different from 3 which was used to generate the data. The same picture is reflected by all the coefficients of models 1 to 3. On the contrary, model 4 has a constant term of 2.94146 which is almost 3. The same is reflected by all the coefficients of model 4.

Table 4, below, is on information criteria. The log likelihood, whose calculation does not take into consideration the number of independent variables in the model favors model 4 where it (the log-likelihood) has value -2804, which is the largest of the four. On the other hand, the AIC, which penalizes a model for its complexity, is against model 4 since it is the one with the smallest value of 3.750, suggesting, again that model 4 is best.

**Table 4: Information Criteria**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Model 1 | Model 2 | Model 3 | Model 4 |
| Log Likelihood | -3227 | -3203. | -2872 | -2804 |
| AIC | 4.308 | 4.279 | 3.836 | 3.750 |

The medians shown in table 5 support the facts given by table 4 that ignoring the structure of the data distorts the efficiency levels of the units under study. Model 4, here gives the largest median. From this table (table 5) the levels of model 3 are very close to those of model 4. Regardless of this, the test proved that the difference between the two models is significant, as will reflected by tests.

**Table 5: Medians of the Models**

|  |  |  |  |
| --- | --- | --- | --- |
| Model 1 (OLS) | Model 2 (RC) | Model 3 (OLS) aggregated | Model 4 (MLM) |
| 0.1937690 | 0.2186613 | 0.2794423 | 0.2804799 |

Because the distribution of the inefficiency levels is not normal, it was not appropriate to use parametric tests to compare the inefficiency levels. Wilcoxon-signed rank test was used to compare the medians of the distributions. First it was model 4 compared to model 1. Since the data generated was too large for the Wilcoxon tables, the normal approximation to Wilcoxon was used. The formula used to calculate the test statistic is:



where *T* = *min*{*T*+*,T*−} as understood from the Wilcoxon test,  and

q

The hypothesis: *H*0 : *η*1 = *η*4 was tested against *H*0 : *η*1 ̸= *η*4, where *η*1 is the median of model 1 and *η*4 is the median of model 4.

The test statistic was found to be −8*.*14, leading to the rejection of *H*0 at the 1% level of significance. This suggests that the medians of two models are different.

The next test compared models 3 and 4. The test statistic here was found to be −2*.*61, which also led to the rejection of *H*0 at the 1% level of significance. A summary of the tests described above is given in table 6 below

**Table 6: Model Comparisons**

|  |  |  |  |
| --- | --- | --- | --- |
| Compared Models | Test Statistic | Critical value | Decision |
| Model 1 vs Model 4 | -8.14 | 2.5758 | Reject *H*0 |
| Model 3 vs Model 4 | -2.61 | 2.5758 | Reject *H*0 |

It is model 4 which is of interest. The median tests here compared efficiency levels from this model to the current commonly used models, the single-level, which completely, ignores the information from levels of the data other than the level of interest and one that ignores the structure but considers information from other levels through disaggregation. Both tests suggest that model 4 has results that are different from those of the other two. The difference between the levels of models 2 and 4, as can be seen from table 4 above, is bigger than that between models 3 and 4 so the significance of this difference is implied.

Researches carried-out so far on efficiency levels produced by different models have dwelt on ranks and not on actual efficiency levels, trying to compare the models. The comparisons have not been statistical since conclusions are drawn from merely looking at the displayed ranks. Findings of this research in as far as ranks are concerned, are similar to those found by researchers like (Lordan, 2009) that the ranks of the different models of interest are correlated. Table 7 displays the rank-order correlation coefficients of the different pairs of models. The greatest correlation-coefficient, from the table, is between model 1 and model 2. The weakest of these correlation coefficients is the one between model 1 and the proposed model, model 4. The size of the correlation-coefficient between model 1 and model 4 suggests that using the fixed coefficients model for clustered data significantly distorts the rankings of the study units.

**Table 7: Spearman’s Rank Correlation-Coefficients of the Models**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Model 1 | Model 2 | Model 3 | Model 4 |
| Model 1 | 1.00000 |  |  |  |
| Model 2 | 0 9876038 | 1.00000 |  |  |
| Model 3 | 0.7757221 | 0.7643373 | 1.00000 |  |
| Model 4 | 0.7548816 | 0.7656512 | 0.9557342 | 1.00000 |

Table 8 shows the variances of the components of the error term. Even these variances support the superiority of model 4. It (model 4) has the smallest variances, implying that it is the most consistent of the 4 models. Model 1 has the largest variances. This means that it is the least consistent of all the four models.

**Table 8: Composite error variances**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Model 1 | Model 2 | Model 3 | Model 4 |
| v | 2.36719 | 1.16173 | 1.46084 | 1.07481 |
| u | 5.59774 | 2.25481 | 3.52314 | 1.91420 |
| *ϵ* | 2.82222 | 2.53649 | 2.23248 | 2.19531 |

The last table here, table 9, shows the values of *λ*. The random coefficients models have higher values of this measure. Although model 4 has the smallest values of both *σu* and *σv*, it has a larger value of *λ*. For all the four models, the greater part of the variation in the composite-error values is due to inefficiency rather than white noise. This is well pronounced in random-coefficients models.

**Table 9: Asymmetry measure **

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Model 1 | Model 2 | Model 3 | Model 4 |
| *λ* | 1.53776 | 1.94092 | 1.55297 | 1.78096 |

# Conclusion

The study aimed at proposing a multilevel stochastic frontier model. It was discovered that a miltilevel model is a special random coefficients model. The (MLM) can be expressed and was expressed in such a way that it is a random coefficients model in everything except that it has variables from two or more levels as well as interactions. It is this form of the (MLM) which was proposed and brought to the frontier modelling environment by assigning to it a two-component error term.

The proposed model was used to establish the effects of ignoring the structure of data in efficiency analysis. Clustered-panel data which composed of 50 clusters, each with 6 subjects, was generated in Rstudio. The data was panel, with five repeated measurements per subject. This resulted in a total of 1500 observations. The data was generated in such a way that there was a panel-level variable, a subject-level variable and a cluster-level variable. Four stochastic frontier models were run at subject level. The four were, two that ignored information from the cluster level, differing only in that one had fixed coefficients while the other had random coefficients, and two that considered information from the cluster level, differing in that one used disaggregation while the other was multilevel.

After running all the four models several times, once with specific normal distributions of x1 and x2, results from the four models show that the proposed model had efficiency levels that were significantly different from each of the other three models. It was evident that the results from the multilevel model were the best because the data were generated from a multilevel model with known parameters and when the frontier models were run in Limdep, it is the multilevel model that had coefficients that were equal to those used in the generation of the dataset.

The proposed model was found to perform better than all the conventional models. Ignoring the structure of data does not only lead to the underestimation of standard errors but also falsifies the efficiency-levels of the subjects. The results found from this simulation study suggest that hierarchical organizations like governments, schools, banks, hospitals and many others have had their efficiency levels underestimated for long and require the use of the proposed model for improved efficiency levels.

**Disclaimer (Artificial intelligence)**

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

# References

1. Aiello F. and Bonann G., (2015) Multilevel empirics for small banks in local markets, Online at http://mpra.ub.uni-muenchen.de/64399/ MPRA Paper No. 64399, posted 17. May 2015 04:54 UTC.
2. Aigner D, Lovell K and Schmidt P. Formulation and Estimation of Stochastic Frontier Models. Journal of Econometrics 1977; 6: 21-37.
3. Chandel, R.b.S.; Khan, A.; Li, X.; Xia, X. Farm-Level Technical Efficiency and Its Determinants of Rice Production in Indo-Gangetic Plains: A Stochastic Frontier Model Approach. Sustainability 2022, 14, 2267. https://doi.org/10.3390/ su14042267
4. El-Horbaty Yahia S. and Hanafy Eman M. (2018) Some Estimation Methods and Their Assessment in Multilevel Models: A Review, Biostatstics and Biometrics open access journal volume issue 3.
5. Galbraith S., Daniel J. A. and Vissel B. A Study of Clustered Data and Approaches to Its Analysis, J. Neurosci., August 11, 2010 30(32):10601–10608 10603.
6. Gelman A. and Hill J. (2007) Data Analysis Using Regression and Multilevel/Hierarchical Models, Cambridge University Press, New York
7. Goldstein, H. (1986). Multilevel mixed linear analysis using iterative generalized least squares. Biometrika, 73, 43–56.
8. Greene, W.H., (2005). Fixed and Random Effects in Stochastic Frontier Models. Journal of Productivity Analysis 23(1/January), 7-32.
9. Harville, D.A. (1977) Maximum Likelihood Approaches to Variance Components Estimation and to Related Problems. Journal of the American Statistical Association, 72, 320 338. http://dx.doi.org/10.1080/01621459.1977.10480998
10. Heck R. H., Thomas S. L., Tabata L. N., (2014) Multilevel and Longitudinal Modelling with IBM SPSS, Routledge, UK.
11. Hung T. Nguyen (2020) A closer look at stochastic frontier analysis in economics, Asian Journal of Economics and Banking Vol. 4 No. 3, 2020 pp. 3-28 Emerald Publishing Limited 2615-9821 DOI 10.1108/AJEB-07-2020-0032.
12. Johnes J. (2006b) Measuring efficiency: a comparison of multilevel modelling and data envelopment analysis in the context of higher education, Blackwell Publishing Ltd, Bulletin of Economic Research 58:2, 2006, 0307–3378.
13. Laird, N. M., and Ware, J. H. (1982). Random-effects models for longitudinal data. Biometrics, 38, 963–974.
14. Liouaeddine M. and Bijou M. (2021) Multilevel modelling an innovative tool for analyzing clustered data: application to education and its projection on environmental management Bijou E3S Web of Conferences 319, 0100 (2021) VIGISAN

2021.

1. Lordan Grace (2009) Considering Endogeneity, Quality of Care and Casemix- A Hierarchical Random Parameters Approach To Measuring Efficiency For Out of Hours Primary Care Services in Ireland. Applied Economics, Taylor Francis (Routledge), 41 (26), pp.3411-3423.
2. Martha A., Giorgio V. and Gianmaria M. (2011) Multilevel and Stochastic Frontier Models: A comparison and a joint approach of their performances when investigating panel data, Universit‘a degli Studi di Milano – Bicocca.
3. McNeish D. M. (2014) Analyzing Clustered Data with OLS Regression: The Effect of a Hierarchical Data Structure, Multiple Linear Regression Viewpoints, 2014, Vol. 40(1)
4. Meeusen, W. and J. van den Broeck (1977), ”Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error” International Economic Review, 18, 435-444.
5. Mundfrom, D., and Schultz, M. (2002). A Monte Carlo simulation comparing parameter estimates from multiple linear regression and hierarchical linear modeling. Multiple Linear Regression Viewpoints, 28, 18-21
6. Patterson, L., and Goldstein, H. (1991). New statistical methods for analyzing social structures: An introduction to multilevel models. British Educational Research Journal, 17, 387–393.
7. Pryseley, A., Tchonlafi, C., Verbeke, G., and Molenberghs, G. (2011). Estimating negative variance components from Gaussian and non-Gaussian data: A mixed models approach. Computational Statistics and Data Analysis, 55(2), 1071–1085. https://doi.org/10.1016/j.csda.2010.09.002
8. Siciliani, L. (2006). Estimating technical efficiency in the hospital sector with panel data: A comparison of parametric and non-parametric techniques. Applied Health Economics and Health Policy, 5(2), 99–116.
9. Swamy, P. A. V.B. (1970). Efficient inference in a random coefficient regression model. Econometrica, Vol. 38, No. 2, pp. 311-323.
10. Thommai Johanna (2019) Simulation Study on the Effects of Ignoring Clustering in Regression Analysis
11. Twisk J. W. R. (2006) Applied Multilevel Analysis. A Practical Guide, Cambridge University Press, New York, United States of America.