Estimating Efficiency Effects in a Panel Data Stochastic Frontier Model

Paul, Satya and Shankar, Sriram

4 June 2018

Estimating Efficiency Effects in a Panel Data Stochastic Frontier Model

Satya Paul[+] and Sriram Shankar[*]

28 September 2018

Abstract

This paper proposes a stochastic frontier model which includes time-invariant unobserved heterogeneity along with the efficiency effects. The efficiency effects are specified by a standard normal cumulative distribution function of exogenous variables which ensures the efficiency scores to lie in a unit interval. The model parameters are consistently estimated by non-linear least squares after removing the individual effects by the usual within transformation. The efficiency scores are directly calculated once the model is estimated. An empirical illustration based on widely used panel data on Indian farmers is presented.

JEL Classification: C51, D24, Q12

Keywords: Fixed effects; Stochastic frontier; Technical efficiency; Standard normal cumulative distribution function; Non-linear least squares.

Acknowledgement: We are grateful to Hung-Jen Wang for providing us the farm level data that we have used for empirical exercise in this paper.

+ Centre for Economics and Governance, Amrita University, India, and 14 Grafton St., Eastlakes, Sydney 2018, Email: satyapaul@outlook.com.

* Corresponding author: Centre for Social Research and Methods, Australian National University, Email: Sriram.Shankar@anu.edu.au.

1. Introduction

There is a vast literature on the measurement of technical (in) efficiency based on stochastic frontier models ever since the pioneering studies of Aigner et al. (1977) and Meeusen van den Broeck (1977). In most of the models, inefficiency is captured by a half normal or truncated normal distribution, and a transformation proposed by Jondrow et al. (1982) (popularly known as JLMS estimator) is utilised to derive the technical inefficiency scores. A number of subsequent stochastic frontier studies have focussed on explaining inefficiency. For this purpose, some studies notably by Kalirajan (1981) and Pitt and Lee (1981) have followed a two-step procedure. In the first step, the production frontier is estimated, and the technical inefficiency scores are obtained for each firm. In the second step, these technical inefficiency scores are regressed against a set of variables which are hypothesized to influence firm's inefficiency. Given the drawbacks associated with the two-step method[1], some recent studies estimate the inefficiency scores and exogenous effects in one single step. Amongst these studies, the most popular are those of Kumbhakar et al. (1991), Huang and Liu (1994) and Battese and Coelli (1995). In order to examine the exogenous influence on inefficiency, these authors parameterize the mean of pre-truncated distribution. These models are further complemented by Caudill and Ford (1993), Caudill et al. (1995) and Hadri (1999) who account for potential heteroscedasticity by parameterizing the variance of pre-truncated distribution. Wang (2002) proposes a more general model that combines these two strands of one-step models.

Availability of Panel data has led to further improvements in the stochastic frontier modelling, allowing for time-invariant unobserved heterogeneity. Some of the earlier panel data stochastic

[1] See, for example, Battese and Coelli (1995), Simar and Wilson (2007) and Wang (2002).

frontier studies treated unobserved heterogeneity as a measure of inefficiency (eg. Schmidt and Sickles, 1984; Kumbhakar, 1990; Battese and Coelli, 1992). This approach does not allow for individual effects (in the traditional sense) to exist alongside inefficiency effects.

Greene (2005) proposes a "true fixed-effect" model, which is essentially a standard fixed-effect panel data model augmented by an additional one-sided error term, whose mean is a function of inefficiency effects. In this model, the heterogeneity is represented by dummy variables and the problem of incidental (nuisance) parameters is encountered. Greene's Monte Carlo simulations reveal that this problem does not affect the frontier coefficients, but it leads to inconsistent variance estimates. A similar result is reported in Wang and Ho (2010). The error variances are important in the stochastic frontier context because they affect the extraction of inefficiency scores from estimated composite residuals (Jondrow et al., 1982).

Chen et al. (2014) and Belotti and Ilardi (2017) adopted different estimation approaches to estimate Greene's model. The estimators proposed in these studies provide consistent estimates of the frontier parameter vector β and composite error variance σ^2 even for small N (number of firms) and T (time observations for each firm). However, these and couple of other studies which explicitly account for 'persistent' (time-invariant) and 'transient' (time varying) inefficiencies, (eg. Colombi et al., 2014) utilise JLMS transformation (Jondrow et al., 1982) to derive the inefficiency scores. As shown in Schmidt and Sickles (1984), the JLMS estimator is not consistent in that the conditional mean or mode of the random variable representing inefficiency component (u) given the composite error (v-u) term, that is, $u|v-u$ never approaches u even when the number of cross-sectional units tends to infinity. However, if the panel data are used, this limitation can be overcome under certain other assumptions, some of which may be less realistic (Parmeter and Kumbhakar, 2014)[2].

Parmeter and Kumbhakar (2014) discuss a distribution free inefficiency effects model which was first proposed in Simar et al. (1994) and later explained in Wang and Schmidt (2002) and Alvarez et al. (2006). Parmeter et al. (2017) non-parametrically estimate distribution free inefficiency effects using a partly linear model initially proposed by Robinson (1988). This model is similar to the one proposed by Deprins and Simar (1989a, 1989b) and extended in Deprins (1989). Paul and Shankar (2018) propose a distribution free efficiency effect model to estimate technical efficiency scores[3]. The efficiency effects are specified by a standard normal cumulative distribution function of exogenous variables which ensures the efficiency scores to lie in a unit interval. Their model eschews one-sided error term present in almost[4] all the existing inefficiency effects models.

However, none of the existing distribution free models including more recent ones by Parmeter et al. (2017) and Paul and Shankar (2018) account for unobserved heterogeneity. The present paper extends Paul and Shankar's (2018) model to account for unobserved heterogeneity within the framework of a panel data stochastic frontier. While this technique can be applied to

---

[2] Battese and Coelli (1988) have proposed an alternative estimator $(E(\exp\{-u\} \mid v-u))$. Kumbhakar and Lovell (2000, pp.77-79) discuss this and the JLMS estimator in details and also refer to related findings of Horrace and Schmidt (1996).

[3] In the efficiency literature, the term 'distribution free' is mentioned in Parmeter and Kumbhakar (2014) to refer to the fact that inefficiency estimation need not utilize the truncated normal distribution. Parmeter and Kumbhakar (2014) utilize a scaling function and Paul and Shankar (2018) use a cumulative distribution function to derive efficiency scores. The relevant details are provided in Section 2 of this paper.

[4] Even though the model as proposed in Parmeter and Kumbhakar (2014) requires no distributional assumptions for the inefficiency term, it does invoke the scaling property in which the inefficiency term is initially assumed to have a basic distribution such as half or truncated normal distribution. Further, Parmeter et al. (2017) make no distributional assumptions concerning the inefficiency term but the estimation is performed in a non-parametric framework

stochastic frontiers of any type, production, cost or any other, the analytical framework and empirical application presented in this paper are specific to a production frontier. The parameters of the production function and efficiency effect specification are estimated by non-linear least squares (NLS) after removing the individual effects by the usual within transformation. Unlike existing parametric stochastic frontier models, the JLMS transformation is not required to compute the efficiency scores. These scores are directly calculated once the model is estimated.

The paper is organised as follows. Section 2 provides a review of existing panel data based stochastic frontier models of inefficiency. In Section 3, we propose a stochastic frontier model to estimate efficiency effects, accounting for time-invariant unobserved heterogeneity. An empirical illustration based on panel data on Indian farmers is presented in Section 4. Section 5 provides conclusions.

2. **A Review of Literature on Efficiency Measurement Based on Panel Data Stochastic Frontier Models**

The literature on efficiency measurement based on panel data stochastic frontier is quite rich and comprehensive. However, our review of literature presented below is brief and selective. It covers topics such as unobserved heterogeneity, true fixed effects, persistent and time varying inefficiencies, and distribution free (in) efficiency effects.

(i) Modelling Unobservable Firm Effects as a Measure of Inefficiency

The role of unobservable individual effects in the panel data estimation of stochastic frontier models has been recognised for long. In some of the early panel data stochastic frontier studies, individual effects are interpreted as inefficiency. For example, Schmidt and Sickles (1984) consider the following stochastic production frontier specification.

$$y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it}, \qquad i = 1, \ldots, N; \; t = 1, \ldots, T. \tag{1}$$

where y_{it} is log of output and x_{it} is a vector whose values are functions of input quantities and time, i and t are cross section and time subscripts respectively, α_i is time-invariant unobserved firm-specific (individual) effect, and ε_{it} is a random noise term. Equation (1) is consistently estimated by 'within group' ordinary least squares. After the model parameters are estimated, individual effects are recovered and then adjusted to conform to an inefficiency interpretation as

$$\hat{\alpha}_i^* = \alpha - \hat{\alpha}_i \qquad \textit{where } \alpha = \max \hat{\alpha}_i \tag{2}$$

That is, inefficiency is measured as the difference between a particular firm's fixed effects and the firm that has the highest estimate of the fixed effects in the sample. By interpreting the firm specific term as 'inefficiency' any unmeasured time invariant cross firm heterogeneity is assumed away. The inefficiency estimates so obtained are time-invariant. Obviously, this approach does not allow for individual effects (in the traditional sense) to exist alongside inefficiency effects.

The time-invariant inefficiency assumption has been relaxed in a number of subsequent studies, including Kumbhakar (1990) and Battese and Coelli (1992). These studies specify inefficiency (u_{it}) as a product of two components. One of the components is a function of time and the other is an individual specific effect so that $u_{it} = G(t) \times u_i$. For example, in Battese and Coelli (1992) $G(t) = \exp\left[-\eta\left(t-T\right)\right]$[5] and $u_i \sim N^+\left(\mu, \sigma^2\right)$[6]. In these models, however, the time varying

[5] η is an unknown scale parameter of the exponential function.

[6] $N^+\left(\mu, \sigma^2\right)$ refers to truncated normal distribution.

pattern of inefficiency is the same for all individuals, so the problem of inseparable inefficiency and individual heterogeneity remains.

(ii) True Fixed Effects Models

Greene (2005) has strongly argued that inefficiency effect and the time- invariant firm-specific effect are different and should be accounted for separately in the estimation. If the firm-specific heterogeneity is not adequately controlled for, then the estimated inefficiency may be picking up firm-specific heterogeneity in addition to or even instead of inefficiency. Thus, the inability of a model to estimate individual effects in addition to the inefficiency effect poses a problem for empirical research. Greene (2005) proposed the following 'True Fixed Effects' (TFE) model which account for unobserved firm specific heterogeneity along with time varying inefficiency.

$$y_{it} = \alpha_i + x_{it}\beta + v_{it} - u_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it} \tag{3}$$

Assuming that the inefficiency term u_{it} is half normally distributed, that is, $u_{it} \sim N^{+}\left(0, \sigma^2\right)$, the log likelihood function for the fixed effects stochastic frontier model is expressed as

$$\log L = \sum_{i=1}^{N}\sum_{t=1}^{T}\log\left[\frac{2}{\sigma}\Phi\left(-\lambda\left(\frac{y_{it}-\alpha_i - x_{it}\beta}{\sigma}\right)\right)\phi\left(\left(\frac{y_{it}-\alpha_i - x_{it}\beta}{\sigma}\right)\right)\right] \tag{4}$$

where $\phi(.)$ and $\Phi(.)$ are the probability and cumulative density functions of a standard normal distribution respectively, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$ is the standard deviation of the composite error term $\varepsilon_{it} = v_{it} - u_{it}$ and $\lambda = \frac{\sigma_u}{\sigma_v}$ is the ratio of inefficiency standard deviation to noise standard deviation. Maximization of the unconditional log likelihood function in (4) is done by 'brute force' even in the presence of possibly thousands of nuisance (incidental) parameters by using

Newton's method. Based on Monte Carlo simulation, Greene shows that β estimates are not biased but the residual estimates are biased possibly due to incidental parameters problem[7].

Wang and Ho (2010) eliminate incidental parameters by either first differencing or within transformation. Their model is specified as:

$$
\begin{aligned}
& y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it} , \\
& \varepsilon_{it} = v_{it} - u_{it} , \\
& v_{it} \sim N\left(0, \sigma_v^2\right), \\
& u_{it} = h_{it} \times u_i^* , \\
& h_{it} = f(z_{it}\delta) , \\
& u_i^* \sim N^+\left(\mu, \sigma_u^2\right).
\end{aligned} \tag{5}
$$

u_{it} is the technical inefficiency and z_{it} is a vector of variables explaining the inefficiency. The model exhibits the "scaling property" in the sense that, conditional on z_{it}, the one-sided error term equals a scaling function h_{it} multiplied by a one-sided error distributed independently of z_{it}. With this property, the shape of the underlying distribution of inefficiency is the same for all individuals, but the scale of the distribution is stretched or shrunk by observation-specific factors z_{it}. The time-invariant specification of u_i^* allows the inefficiency u_{it} to be correlated over time for a given individual. On first differencing, the above equations result in the following:

7 The incidental parameters problem is first defined in Neyman and Scott (1948) and surveyed in Lancaster (2000).

$$\Delta\tilde{y}_i = \Delta\tilde{x}_i\beta + \Delta\tilde{\varepsilon}_i,$$

$$\Delta\tilde{\varepsilon}_i = \Delta\tilde{v}_i - \Delta\tilde{u}_i,$$

$$\Delta\tilde{v}_i \sim MVN\left(0,\Sigma\right), \tag{6}$$

$$\Delta\tilde{u}_i = \Delta h_{it} \times u_i^*,$$

$$u_i^* \sim N^+\left(\mu,\sigma_u^2\right)$$

where $\Delta\tilde{w}_i = \left(\Delta w_{i2},\Delta w_{i3},...,\Delta w_{iT}\right)'$, $w \in \{y,x,\varepsilon,u,v\}$. The first-difference introduces correlations of Δv_{it} within the ith panel, and the $\left(T-1\right)\times\left(T-1\right)$ variance-covariance matrix of the multivariate normal distribution (MVN) of $\Delta\tilde{v}_i$ is given by

$$\Sigma = \sigma_v^2 \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \ddots & \ddots & \ddots & \ddots & -1 \\ 0 & 0 & \dots & -1 & 2 \end{bmatrix} \tag{7}$$

Marginal likelihood function is then derived and estimation is performed by numerically maximising the marginal log-likelihood function of the model (see Wang and Ho, 2010, p. 288 for details). Monte Carlo simulations carried out in their paper indicate that while the incidental parameters problem does not affect the estimation of slope coefficients, it does introduce bias in the estimated model residuals. Since the inefficiency estimation is based on residuals, incident parameter problem should be of concern to empirical researchers, particularly when T is not large.[8]

[8] Wang and Ho (2010) also estimated their model after within transformation and the results of Monte Carlo simulations do not alter the conclusions.

Chen et al. (2014) suggest an alternative to the TFE treatment of Wang and Ho (2010). Specifically, they propose a consistent marginal maximum likelihood estimator (MMLE) for the TFE model exploiting a within-group data transformation and the properties of the closed skew normal (CSN) class of distributions (Gonzalez-Farias et al., 2004). They also conduct a simulation exercise and do not encounter any bias in the estimation of variance that Greene (2005) and Wang and Ho (2010) have found in their studies.

Belotti and Ilardi (2018) propose two alternative consistent estimators which extend the Chen et al. (2014) results in different directions. The first estimator is a marginal maximum simulated likelihood estimator (MMSLE) that can be used to estimate both homoscedastic and heteroskedastic normal-half normal and normal-exponential models. This estimator allows only the time-invariant inefficiency effects. The second is a U-estimator based on all pairwise quasi-likelihood contributions constructed exploiting the analytical expression available for the marginal likelihood function when T = 2. This strategy allows them to provide a computationally feasible approach to estimate normal-half normal, normal-exponential and normal-truncated normal models in which inefficiencies can be heteroskedastic and may follow a first-order autoregressive process. This estimator allows the modelling of inefficiency variance[9] as a function of exogenous effects. Finally, the finite sample properties of the proposed estimators are investigated by conducting Monte Carlo simulations. The results show good finite sample properties, especially in small samples.

[9] Existing effects models parameterize the mean of the pre-truncated distribution as a way to study the exogenous influence on inefficiency.

In another related research, Wikstrom (2015) suggests a class of consistent method of moment estimators that goes beyond the normal half-normal TFE model proposed by Greene (2005). This is demonstrated by deriving a consistent normal-gamma TFE estimator.

(iii) Models with Persistent and Time Varying Inefficiencies

In some panel data-based models, technical inefficiency is viewed as consisting of two components, namely, persistent (long run) inefficiency and time varying (short run) inefficiency. The persistent inefficiency is time-invariant and could arise due to the presence of rigidity within a firm's organisation and production process. Unless there is a change in something that affects management practices at the firm (for example, new government regulations or a change in ownership), it is unlikely that persistent inefficiency will change. The time varying inefficiency could be due to non-organisational factors that can be reduced/ removed in the short run even in the presence of organisational rigidities[10]. The models proposed by Kumbhakar (1991), Kumbhakar and Heshmati (1995) and Kumbhakar and Hjalmarsson (1993, 1995) treat firm effects as persistent inefficiency and include another component to capture time varying technical inefficiency and thus do not account for the heterogeneity effects. The task of estimating these two inefficiencies while also allowing for firm-effects heterogeneity is undertaken in Tsionas and Kumbhakar (2012) and Colombi et al. (2014). The model proposed by these authors can be written as (see Kumbhakar et al., 2012):

[10] Colombi et al. (2014) have clarified the difference between persistent and time-varying inefficiencies by giving an example of a hospital which has more capacity (beds) than the optimal required level, but downsizing may be a long-run process due to social pressure. This implies that the hospital has a long-run inefficiency since this gap cannot be completely recovered in the short-run. But this hospital may increase its efficiency in the short-run by reallocating the work force across different activities. Thus, some of the physicians' and nurses' daily working hours might be changed to include other hospital activities such as acute discharges. This is a short-run improvement in efficiency. Hence, the hospital continues to suffer from long run inefficiency due to excess capacity, but the time varying activities have improved part of its short-run inefficiency.

$$
\begin{aligned}
&y_{it} = \alpha + w_i + x_{it}\beta + v_{it} - u_{it} - h_i \\
&w_i \sim N^+\left(0, \sigma_w^2\right) \\
&v_{it} \sim N\left(0, \sigma_v^2\right) \\
&u_{it} \sim N^+\left(0, \sigma_u^2\right) \\
&h_i \sim N^+\left(0, \sigma_h^2\right)
\end{aligned}
\tag{8}
$$

In this model w_i, u_{it}, h_i represent respectively firm-specific unobserved heterogeneity, transient inefficiency and persistent inefficiency. Fillipini and Greene (2016) develop a practical full information maximum simulated likelihood estimator for this model in order to reduce the extreme complexity of the log likelihood function in Colombi et al. (2014).

(iv) Distribution Free Models of (In)efficiency Measurement

Parmeter and Kumbhakar (2014) discuss a model possessing the scaling property which can be estimated without making any distributional assumption. Their model can be written as

$$y_{it} = x_{it}\beta + v_{it} - g\left(z_{it}\gamma\right)u_{it} \tag{9}$$

where $g\left(z_{it}\gamma\right) = e^{z_{it}\gamma}$ is the scaling function and u_{it} the basic distribution such as half-normal or truncated normal. The conditional mean of *y*, given *x* and *z*, is

$$E\left(y_{it} \middle| x_{it}, z_{it}\right) = x_{it}\beta - e^{z_{it}\gamma}\mu \tag{10}$$

where $\mu = E\left(u_{it}\right)$. The equation (9) can be re-written as

$$y_{it} = x_{it}\beta - e^{z_{it}\gamma}\mu + v_{it} - e^{z_{it}\gamma}\left(u_{it} - \mu\right) = x_{it}\beta - e^{z_{it}\gamma}\mu + \varepsilon_{it} \tag{11}$$

where $\varepsilon_{it} = v_{it} - e^{z_{it}\gamma}\left(u_{it} - \mu\right)$ is independent but not identically distributed. This model can be estimated with nonlinear least squares by minimizing $\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it} - x_{it}\beta + e^{z_{it}\gamma}\mu\right)^2$.

Parmeter et al. (2017) estimate the following partly linear regression model initially proposed by Robinson (1988), which does not invoke the scaling property.

$$y_{it} = x_{it}\beta + v_{it} - u_{it} = x_{it}\beta - g\left(z_{it}\gamma\right) + v_{it} - \left(u_{it} - g\left(z_{it}\gamma\right)\right) = x_{it}\beta - g\left(z_{it}\gamma\right) + \varepsilon_{it} \qquad (12)$$

where $\varepsilon_{it} = v_{it} - \left(u_{it} - g\left(z_{it}\gamma\right)\right)$ and $E\left(u_{it}\right) = g\left(z_{it}\gamma\right) > 0$. To estimate β, the following equation is required.

$$y_{it} - E\left(y_{it} \middle| z_{it}\right) = \left(x_{it} - E\left(x_{it} \middle| z_{it}\right)\right)\beta + \varepsilon_{it} \qquad (13)$$

Since, $E\left(y_{it} \middle| z_{it}\right)$ and $E\left(x_{it} \middle| z_{it}\right)$ are unknown, to obtain consistent estimate of β for the partly linear model of Robinson (1988) the conditional means are replaced with their nonparametric estimates.

As pointed out in Parmeter and Kumbhakar (2014), the above two models, (11) and (12), suffer from certain limitations. First, to avoid identification issues, z cannot contain a constant term in models (11) and (12). Second, in model (11), since ε depends on z through $e^{z\gamma}$, x and z cannot contain common elements. However, Parmeter et al. (2017) show that $x - E\left(x \middle| z\right)$ in (13) is uncorrelated with ε and hence the correlation between z and x is not an issue. Finally, it is possible to obtain negative estimates of $g\left(z\right)$ in model (12) which is inconsistent with the notion that $g\left(z\right)$ represents average inefficiency.

Paul and Shankar (2018) propose a distribution free model wherein the efficiency effects are specified by a standard normal cumulative distribution function of exogenous variables. This ensures the efficiency scores to lie in a unit interval. Their model eschews one-sided error term present in almost all the existing inefficiency effects models. The model contains only a statistical noise term (v), and its estimation is done in a straightforward manner using the non-linear least squares. Once the parameters are estimated, the efficiency scores are calculated directly.

However, all the existing distribution free models including more recent ones by Parmeter et al. (2017) and Paul and Shankar (2018) do not account for unobserved heterogeneity. In the next section, we extend Paul and Shankar's (2018) stochastic frontier model to account for unobserved heterogeneity.

3. The Model

3.1 Model Specification

We propose the following TFE stochastic production frontier efficiency effects model which accounts for time-invariant unobserved heterogeneity.

$$Y_{it} = \exp(\alpha_i + x_{it}\beta + v_{it} + \frac{1}{\mu}\ln\left[H(z_{it}\gamma)\right]u_{it}) \quad i = 1,...,N; t = 1,...,T_i \tag{14}$$

where $\frac{1}{\mu}\ln\left[H(z_{it}\gamma)\right]u_{it}$ is the scaled one-sided error term, Y_{it} is the quantity of output, x_{it} is a vector whose values are functions of input quantities and time, and β is the corresponding coefficient vector $(K_1 \times 1)$. α_i is firm-specific unobserved effect, and v_{it} represents the random noise. $H(z_{it}\gamma)$ represents technical efficiency and is required to lie between 0 and 1, that is, $0 \le H(z_{it}\gamma) \le 1$. Any cumulative distribution function (cdf) will satisfy this property. We assume the efficiency term to take a probit functional form, that is, $H(z_{it}\gamma)=\Phi(z_{it}\gamma)$, where Φ is a standard normal cdf, z_{it} is a vector containing a constant 1 and exogenous variables[11] assumed to influence efficiency and γ is the corresponding $(K_2\times 1)$ coefficient vector[12]. Equation (14) is written assuming that the panel data are unbalanced. However, in the case of balanced data, T_i is to be replaced by T for all i.

Taking logarithm on both sides of (14), we have

$$y_{it}=\ln(Y_{it})=\alpha_i+x_{it}\beta+\frac{1}{\mu}\ln\left[\Phi(z_{it}\gamma)\right]u_{it})+v_{it}\quad i=1,...,N;t=1,...,T_i \tag{15}$$

This can be re-written as

$$y_{it}=\alpha_i+x_{it}\beta+\ln\left[\Phi(z_{it}\gamma)\right]+\varepsilon_{it} \tag{16}$$

where $\varepsilon_{it}=v_{it}+\frac{1}{\mu}\ln\left[\Phi(z_{it}\gamma)\right]\left(u_{it}-\mu\right)$ is independent but not identically distributed and $\mu=E\left(u_{it}\right)$.

3.2 Within Transformation

The within transformation will eliminate unobserved firm-specific effects (α_i). Thus, on subtracting time averages of the concerned variables, we have

---

[11] A potential limitation of our specification as well most other distribution free inefficiency effects models including the recent one by Parmeter et al. (2017) is that the firms with the same z have the same efficiency. However, in most practical applications if sufficient number of variables are included into the (in) efficiency effects model then it is less likely that any two firms in the same time period or the same firm across different time periods will have the same z vector.

[12] We could have chosen any other function which is not a cumulative distribution function as long as this function is constrained to lie between 0 and 1. For example, we could have chosen $H(z_{it}\gamma)=\frac{1}{1+z_{it}\gamma}$ and restricted $z_{it}\gamma\geq 0$. Another example of a function which is also not a distribution function but whose range lies is the unit interval, is a Gompertz function of the form $G(z_{it}\gamma)=e^{-e^{z_{it}\gamma}}$ (see Simar et al., 1994)). Instead, we chose probit function because it is quite popular in econometric literature and we do not have to impose any constraints on the parameter vector γ so that $0<\Phi(z_{it}\gamma)<1$.

$$y_{it}-\frac{1}{T_i}\sum_{p=1}^{T_i}y_{ip}=\left(x_{it}-\frac{1}{T_i}\sum_{p=1}^{T_i}x_{ip}\right)\beta+\ln\left(\Phi(z_{it}\gamma)\right)-\frac{1}{T_i}\sum_{p=1}^{T_i}\ln\left(\Phi(z_{ip}\gamma)\right)+\varepsilon_{it}-\frac{1}{T_i}\sum_{p=1}^{T_i}\varepsilon_{ip}$$

or

$$y_{it}-\overline{y}_i=\left(x_{it}-\overline{x}_i\right)\beta+\ln\left(\frac{\Phi(z_{it}\gamma)}{\left(\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right)^{1/T_i}}\right)+\varepsilon_{it}-\overline{\varepsilon}_i \tag{17}$$

$$\tilde{y}_{it}=\tilde{x}_{it}\beta+\ln\left(\frac{\Phi(z_{it}\gamma)}{\left(\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right)^{1/T_i}}\right)+\tilde{\varepsilon}_{it}\quad i=1,...,N;t=1,...,T_i$$

where $\overline{w}_i=\frac{1}{T_i}\sum_{p=1}^{T_i}w_{ip}$ and $\tilde{w}_i=w_{it}-\overline{w}_i,\ w\in\left\{y,x,\varepsilon\right\}$.

Equation (17) can be estimated by minimizing the following sum of squared errors with respect to parameter vector θ :

$$Q_{NT}\left(\theta\right)=\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(\tilde{y}_{it}-\tilde{x}_{it}\beta-\ln\left(\frac{\Phi(z_{it}\gamma)}{\left(\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right)^{1/T_i}}\right)\right)^2 \tag{18}$$

where $\theta=\left(\beta',\gamma'\right)'$ is $\left(K\times1\right)$ parameter vector. Equation (17) can be estimated using the nonlinear least squares (NLS).

Equation (17) can be re-written as

$$\tilde{y}_{it}=g\left(\tilde{x}_{it},z_{it},\theta\right)+\tilde{\varepsilon}_{it}=g_{it}+\tilde{\varepsilon}_{it}\quad i=1,...,N;t=1,...,T_i \tag{19}$$

where $g\left(\tilde{x}_{it}, z_{it}, \theta\right) = \tilde{x}_{it}\beta + \ln\left(\frac{\Phi(z_{it}\gamma)}{\left(\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right)^{1/T_i}}\right)$.

Stacking all the time series observations of a firm together equation (19) can be compactly written in matrix notion as

$$\tilde{\mathbf{y}} = \mathbf{g} + \tilde{\boldsymbol{\varepsilon}} \tag{20}$$

where

$$\tilde{\mathbf{y}} = \left(\tilde{y}_{11}, \tilde{y}_{12}, \cdots, \tilde{y}_{1T_1}, \cdots, \tilde{y}_{i1}, \tilde{y}_{i2}, \cdots, \tilde{y}_{iT_i}, \cdots, \tilde{y}_{N1}, \tilde{y}_{N2}, \cdots, \tilde{y}_{NT_N}\right)'$$

$$\mathbf{g} = \{g(\tilde{x}_{11}, z_{11}, \theta), g(\tilde{x}_{12}, z_{12}, \theta), \cdots, g(\tilde{x}_{1T_1}, z_{1T_1}, \theta), \cdots, g(\tilde{x}_{i1}, z_{i1}, \theta), g(\tilde{x}_{i2}, z_{i2}, \theta), \cdots, g(\tilde{x}_{iT_i}, z_{iT_i}, \theta), \cdots, g(\tilde{x}_{N1}, z_{N1}, \theta), g(\tilde{x}_{N2}, z_{N2}, \theta), \cdots, g(\tilde{x}_{NT_N}, z_{NT_N}, \theta)\}'$$

$$\tilde{\boldsymbol{\varepsilon}} = \left(\tilde{\varepsilon}_{11}, \tilde{\varepsilon}_{12}, \cdots, \tilde{\varepsilon}_{1T_1}, \cdots, \tilde{\varepsilon}_{i1}, \tilde{\varepsilon}_{i2}, \cdots, \tilde{\varepsilon}_{iT_i}, \cdots, \tilde{\varepsilon}_{N1}, \tilde{\varepsilon}_{N2}, \cdots, \tilde{\varepsilon}_{NT_N}\right)'$$

Each of these three vectors have a dimension of $\left(L \times 1\right)$, where $L = \sum_{i=1}^{N} T_i$.

Thus equation (18) can be written as

$$Q_L\left(\theta\right) = \sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(\tilde{y}_{it} - g\left(\tilde{x}_{it}, z_{it}, \theta\right)\right)^2 = \tilde{\boldsymbol{\varepsilon}}'\tilde{\boldsymbol{\varepsilon}} = \left(\tilde{\mathbf{y}} - \mathbf{g}\right)'\left(\tilde{\mathbf{y}} - \mathbf{g}\right) \tag{21}$$

Given this setup, Theorems 1 and 2 state the assumptions required for the proposed nonlinear least squares estimator to be consistent.

Theorem 1:

Consider the following conditions.

(i) The parameter space Θ is an open subset of R^K ,

(ii) $Q_L\left(\theta\right)$ is a measurable function of data for all $\theta \in \Theta$ and $Q_L\left(\theta\right)$ is continuous in $\theta \in \Theta$.

(iii) The objective function $Q_L(\theta)$ converges in probability to a non-stochastic function $Q_0(\theta)$, and $Q_0(\theta)$ attains a global maximum at θ_0.

If these conditions hold, then the estimator $\hat{\theta}_{NLS} = \arg\max_{\theta\in\Theta} Q_L(\theta)$ is consistent for θ_0, that is, $\hat{\theta}_{NLS} \xrightarrow{p} \theta_0$. Here, uniform convergence in probability of $Q_L(\theta)$ to $Q_0(\theta) = \text{plim}\, Q_L(\theta)$ in condition (iii) implies that $\text{Sup}_{\theta\in\Theta} \left|Q_L(\theta) - Q_0(\theta)\right| \xrightarrow{p} 0$.

Proof: See Theorem 4.1.1 and related proof in Amemiya (1985).

In case of local maximum, the first derivative of $Q_L(\theta)$ need to exist, but one needs to then focus on the behaviour of $Q_L(\theta)$ and its derivative in the neighbourhood of θ_0.

Theorem 2:

Consider the following conditions.

(i) The parameter space Θ is an open subset of R^K.

(ii) $Q_L(\theta)$ is a measurable function of data for all $\theta \in \Theta$, and $\frac{\partial Q_L(\theta)}{\partial \theta}$ exists and is continuous in an open neighbourhood of θ_0.

(iii) The objective function $Q_L(\theta)$ converges uniformly in probability to $Q_0(\theta)$ in open neighbourhood of θ_0, and $Q_0(\theta)$ attains a local maximum at θ_0.

If these conditions hold, then one of the solutions to $\frac{\partial Q_L(\theta)}{\partial \theta} = \mathbf{0}$ is consistent for θ_0.

Proof: See Theorem 4.1.2 and related proof in Amemiya (1985).

Condition (i) in Theorem 1 allows a global maximum to be on the boundary of the parameter space, whereas in Theorem 2 a local maximum has to be in the interior of the parameter space.

Condition (ii) in Theorem 2 also means continuity of $Q_L(\theta)$ in an open neighbourhood of θ_0, where a neighbourhood $N(\theta_0)$ of θ_0 is open if and only if there exists a ball with centre θ_0 entirely contained in $N(\theta_0)$. Condition (iii) is important in both the Theorems as the maximum, global or local, of $Q_L(\theta)$ must occur at $\theta = \theta_0$. In Theorem 1, (iii) provides the identification condition that θ is unique. For local maximum, if there is only one local maximum then analysis is straight forward as $\hat{\theta}$ is uniquely defined by $\left.\frac{\partial Q_L(\theta)}{\partial \theta}\right|_{\hat{\theta}_{NLS}} = \mathbf{0}$. When there is more than one local maximum then Theorem 2 implies that one of the local maximum is consistent. In such cases it is best to consider the global maximum and apply Theorem 1. For a discussion on this, see Newey and McFadden (1994, p.2117).

The following Theorem establishes the asymptotic distribution for the proposed nonlinear least squares estimator.

Theorem 3:

Consider the following specifications.

(i) The model is given by $\tilde{y}_{it} = g\left(\tilde{x}_{it}, z_{it}, \theta\right) + \tilde{\varepsilon}_{it} = g_{it} + \tilde{\varepsilon}_{it}$

(ii) In the DGP, $E\left(\tilde{\varepsilon}_{it} | x_{it}, z_{it}\right) = 0$ and $E\left(\tilde{\boldsymbol{\varepsilon}}\tilde{\boldsymbol{\varepsilon}}' | x_{it}, z_{it}\right) = \Sigma$, where $\Sigma_{ij} = \sigma_{ij}$.

(iii) The mean function $g(\cdot)$ satisfies $g\left(x, z, \theta^{(1)}\right) = g\left(x, z, \theta^{(2)}\right)$ if and only if $\theta^{(1)} = \theta^{(2)}$

(iv) The matrix

$$\mathrm{A}_0 = \operatorname{plim} \frac{1}{L} \frac{\partial \mathbf{g}'}{\partial \theta} \frac{\partial \mathbf{g}}{\partial \theta'}\bigg|_{\theta_0} \tag{22}$$

exists and is finite non-singular for any sequence $\tilde{\theta}$ such that $\tilde{\theta} \xrightarrow{p} \theta_0$.

(v) $L^{-1/2} \operatorname{plim} \frac{1}{L} \sum_{i=1}^{N} \sum_{j=1}^{T_i} \frac{\partial g_{it}}{\partial \theta} \tilde{\varepsilon}_{it}\bigg|_{\theta_0} \xrightarrow{d} N\left[\mathbf{0}, \mathrm{B}_0\right]$, where

$$\mathrm{B}_0 = \operatorname{plim} \frac{1}{L} \frac{\partial \mathbf{g}'}{\partial \theta} \Sigma \frac{\partial \mathbf{g}}{\partial \theta'}\bigg|_{\theta_0} . \tag{23}$$

Then, the NLS estimator $\hat{\theta}_{NLS}$, defined to be root of the first order conditions $\partial L^{-1} Q_L(\theta) / \partial \theta = \mathbf{0}$ is consistent for θ_0 and

$$L^{-1/2}\left(\hat{\theta}_{NLS} - \theta_0\right) \xrightarrow{d} N\left[\mathbf{0}, \mathrm{A}_0^{-1} \mathrm{B}_0 \mathrm{A}_0^{-1}\right]. \tag{24}$$

Proof: See Appendix A

Specifications (i) to (iii) are needed to establish consistency of $\hat{\theta}_{NLS}$, that is, $\hat{\theta}_{NLS} \xrightarrow{p} \theta_0$ and they also imply that the regression function is correctly specified and regressors are uncorrelated with the errors and that θ_0 is identified. Since probit is a strictly increasing function of its parameters, (iii) is satisfied in Theorem 3. The probability limits in (22) and (23) are with respect to the DGP of *x* and *z*; they become regular limits if *x* and *z* are non-stochastic. Since, $L = \sum_{i=1}^{N} T_i$, consistency of the NLS estimator is obtained by $N \to \infty$ or $T_i \to \infty \; \forall i$.

Given Theorem 3, the resulting asymptotic distribution of the NLS estimator can be expressed as

$$\hat{\theta}_{NLS} \overset{a}{\sim} N\left[\theta, \left(D'D\right)^{-1} D'\Sigma D \left(D'D\right)^{-1}\right] \tag{25}$$

where the derivative matrix $D = \frac{\partial \mathbf{g}}{\partial \theta'}\Big|_{\theta_0}$ and $\frac{\partial \mathbf{g}'}{\partial \theta} = \begin{pmatrix} \frac{\partial g_1}{\partial \theta_1} & \cdots & \frac{\partial g_L}{\partial \theta_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial \theta_K} & \cdots & \frac{\partial g_L}{\partial \theta_K} \end{pmatrix}$. We assume that the law of large numbers applies, so that the plim operator in the definitions of A_0 and B_0 are replaced by limE, and then the limit can be dropped.

3.3 Structure of the Covariance Matrix

Theorem 4 below specifies the error covariance matrix in (19).

Theorem 4:

Consider the following assumptions.

(i) $E\left(v_{it}^2\right)=\sigma_v^2$

(ii) $E\left(v_{it}v_{is}\right)=0 \; if \;\; t \neq s$

(iii) $E\left(v_{it}v_{jt}\right)=0 \; if \;\; i \neq j$

(iv) $E\left(u_{it}\right)=\mu$

(v) $E\left(u_{it}u_{js}\right)=E\left(u_{it}\right)E\left(u_{js}\right)=\mu^2 \;\; \forall \, i \neq j$ and/or $t \neq s$

(vi) $E\left(\left[u_{it}-E\left(u_{it}\right)\right]^2\right)=\sigma_u^2$

(vii) $E\left(v_{it}u_{js}\right)=E\left(v_{it}\right)E\left(u_{js}\right)=0 \;\; \forall \, i, j, t$ and s

If these assumptions hold, then, $E\left(\tilde{\boldsymbol{\varepsilon}}\tilde{\boldsymbol{\varepsilon}}'\right)=\Sigma=\begin{pmatrix} \Sigma_1 & 0 & \cdots & 0 & 0 \\ 0 & \vdots & \vdots & 0 & 0 \\ 0 & 0 & \Sigma_i & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \Sigma_N \end{pmatrix}$ has a dimension of $\left(\sum_{i=1}^{N} T_i \times \sum_{i=1}^{N} T_i\right)_i$. Σ_i has a dimension of $\left(T_i \times T_i\right)$ and is specified as

$$\Sigma_i = \Sigma_{Ai} + \Sigma_{Bi} + \Sigma_{Ci} \tag{26}$$

where

$$\Sigma_{Ai} = \sigma_v^2 \begin{pmatrix} \frac{T_i - 1}{T_i} & \frac{-1}{T_i} & \cdots & \frac{-1}{T_i} \\ \frac{-1}{T_i} & \frac{T_i - 1}{T_i} & \cdots & \frac{-1}{T_i} \\ \vdots & \vdots & & \vdots \\ \frac{-1}{T_i} & & \cdots & \frac{T_i - 1}{T_i} \end{pmatrix},$$

$$\Sigma_{Bi} =$$

$$\frac{\sigma_u^2}{\mu^2} \begin{pmatrix} \{\ln(\Phi(z_{i1}\gamma))\}^2 \left(1 - \frac{2}{T_i}\right) & -\frac{1}{T_i}\left(\left[\ln(\Phi(z_{i1}\gamma))\right]^2 + \left[\ln(\Phi(z_{i2}\gamma))\right]^2\right) & \cdots & -\frac{1}{T_i}\left(\left[\ln(\Phi(z_{i1}\gamma))\right]^2 + \left[\ln(\Phi(z_{iT_i}\gamma))\right]^2\right) \\ -\frac{1}{T_i}\left(\left[\ln(\Phi(z_{i2}\gamma))\right]^2 + \left[\ln(\Phi(z_{i1}\gamma))\right]^2\right) & \{\ln(\Phi(z_{i2}\gamma))\}^2 \left(1 - \frac{2}{T_i}\right) & \cdots & -\frac{1}{T_i}\left(\left[\ln(\Phi(z_{i2}\gamma))\right]^2 + \left[\ln(\Phi(z_{iT_i}\gamma))\right]^2\right) \\ \vdots & \vdots & & \vdots \\ -\frac{1}{T_i}\left(\left[\ln(\Phi(z_{iT_i}\gamma))\right]^2 + \left[\ln(\Phi(z_{i1}\gamma))\right]^2\right) & -\frac{1}{T_i}\left(\left[\ln(\Phi(z_{iT_i}\gamma))\right]^2 + \left[\ln(\Phi(z_{i2}\gamma))\right]^2\right) & \cdots & \{\ln(\Phi(z_{iT_i}\gamma))\}^2 \left(1 - \frac{2}{T_i}\right) \end{pmatrix}$$

$\Sigma_{Ci} = \frac{\sigma_u^2}{\mu^2 T_i^2} \sum_{k=1}^{T_i} \{\ln(\Phi(z_{ik}\gamma))\}^2 \iota_i \iota_i'$, where ι_i is a $(T_i \times 1)$ vector of 1's.

Proof: See Appendix B

3.4. Recovering the Individual Fixed Effects

Even though the individual effects α_i's are not estimated in the model, their values can be retrieved as follows:

$$\hat{\alpha}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} \left(y_{it} - x_{it}\hat{\beta} - \ln(\Phi(z_{it}\hat{\gamma})) \right) \quad i = 1, \ldots, N \tag{27}$$

The estimates of individual fixed effects thus obtained would be consistent if $T_i \to \infty$. The hat symbol on the right-hand side of the above equation refers to values estimated with nonlinear least squares.

3.5 Estimating $\hat{\sigma}_v^2$

Since σ_v^2 is unknown, in order to obtains its estimates we have to assume a distribution for the term u_{it} in (15). We assume it to be Gamma distributed, that is, $u_{it} \sim \Gamma(\mu\delta, \delta)$. We then have

$$E(u_{it}) = \frac{\mu\delta}{\delta} = \mu \tag{28}$$

$$\sigma_u^2 = V(u_{it}) = E\left(\left[u_{it} - E(u_{it})\right]^2\right) = \frac{\mu\delta}{\delta^2} = \frac{\mu}{\delta} \text{ and} \tag{29}$$

$$E\left(\left[u_{it} - E(u_{it})\right]^3\right) = \frac{2\mu\delta}{\delta^3} = \frac{2\mu}{\delta^2} \tag{30}$$

These properties in (16) imply that

$$E(\varepsilon_{it}) = 0 \tag{31}$$

$$\begin{aligned} E\left(\varepsilon_{it}^2\right) &= \sigma_v^2 + \frac{1}{\mu^2}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^2 V(u_{it}) = \sigma_v^2 + \frac{1}{\mu^2}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^2 \frac{\mu}{\delta} \\ &= \sigma_v^2 + \frac{1}{\mu\delta}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^2 \end{aligned} \tag{32}$$

$$\begin{aligned} E\left(\varepsilon_{it}^3\right) &= \frac{1}{\mu^3}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^3 E\left(\left[u_{it} - E(u_{it})\right]^3\right) = \frac{1}{\mu^3}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^3 \frac{2\mu}{\delta^2} \\ &= \frac{2}{(\mu\delta)^2}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^3 \end{aligned} \tag{33}$$

A consistent estimate of σ_v^2 can now be obtained[13] from the residuals in (15) by means of

[13] Since $\hat{\varepsilon}_{it}^3 = E\left(\varepsilon_{it}^3\right) + e_{it} = \frac{2}{(\mu\delta)^2}\left[\ln\left(\Phi(z_{it}\hat{\gamma})\right)\right]^3 + e_{it}$, $\frac{2}{\left(\widehat{\mu\delta}\right)^2}$ is the coefficient estimate of this regression. Here e_{it} is the random error term.

$$\frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}\hat{\varepsilon}_{it}^3\left[\ln\left(\Phi(z_{it}\hat{\gamma})\right)\right]^3}{\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\ln\left(\Phi(z_{it}\hat{\gamma})\right)\right]^6} = \frac{2}{\left(\widehat{\mu\delta}\right)^2} \Rightarrow \widehat{\mu\delta} = \sqrt{\frac{2\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\ln\left(\Phi(z_{it}\hat{\gamma})\right)\right]^6}{\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[\hat{\varepsilon}_{it}\ln\left(\Phi(z_{it}\hat{\gamma})\right)\right]^3}} \tag{34}$$

and

$$\hat{\sigma}_v^2 = \frac{1}{L-K}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(\hat{\varepsilon}_{it}^2 - \frac{\left[\ln\left(\Phi(z_{it}\hat{\gamma})\right)\right]^2}{\widehat{\mu\delta}}\right) \tag{35}$$

Hence, by substituting $\hat{\sigma}_v^2$ from the above equation in (26) the feasible variance-covariance[14] matrix is given[15] by

$$\hat{\Sigma} = \begin{pmatrix} \hat{\Sigma}_1 & 0 & \cdots & 0 & 0 \\ 0 & \vdots & \vdots & 0 & 0 \\ 0 & 0 & \hat{\Sigma}_i & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \hat{\Sigma}_N \end{pmatrix} \tag{36}$$

where

$$\hat{\Sigma}_i = \hat{\Sigma}_{Ai} + \hat{\Sigma}_{Bi} + \hat{\Sigma}_{Ci} \tag{37}$$

[14] From (28) and (29) we get $\frac{\sigma_u^2}{\mu^2} = \frac{1}{\mu\delta}$ and this is used in the variance-covariance matrix below.

[15] Since $\hat{\varepsilon}_{it}^2 = \sigma_v^2 + \frac{1}{\widehat{\mu\delta}}\left[\ln\left(\Phi(z_{it}\hat{\gamma})\right)\right]^2 + e_{it}$, $\hat{\sigma}_v^2$ is obtained by regressing $\hat{\varepsilon}_{it}^2 - \frac{1}{\widehat{\mu\delta}}\left[\ln\left(\Phi(z_{it}\hat{\gamma})\right)\right]^2$ on a constant. Here e_{it} is the random error term.

$$\hat{\Sigma}_{Ai} = \hat{\sigma}_v^2 \begin{pmatrix} \frac{T_i - 1}{T_i} & \frac{-1}{T_i} & \cdots & \frac{-1}{T_i} \\ \frac{-1}{T_i} & \frac{T_i - 1}{T_i} & \cdots & \frac{-1}{T_i} \\ \vdots & \vdots & & \vdots \\ \frac{-1}{T_i} & \cdots & & \frac{T_i - 1}{T_i} \end{pmatrix},$$

$\Sigma_{Bi} =$

$$\frac{1}{\left(\widehat{\mu\delta}\right)} \begin{pmatrix} \left\{\ln\left(\Phi(z_{i1}\hat{\gamma})\right)\right\}^2\left(1-\frac{2}{T_i}\right) & -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{i1}\hat{\gamma})\right)\right]^2+\left[\ln\left(\Phi(z_{i2}\hat{\gamma})\right)\right]^2\right) & \cdots & -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{i1}\hat{\gamma})\right)\right]^2+\left[\ln\left(\Phi(z_{iT_i}\hat{\gamma})\right)\right]^2\right) \\ -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{i2}\hat{\gamma})\right)\right]^2+\left[\ln\left(\Phi(z_{i1}\hat{\gamma})\right)\right]^2\right) & \left\{\ln\left(\Phi(z_{i2}\hat{\gamma})\right)\right\}^2\left(1-\frac{2}{T_i}\right) & \cdots & -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{i2}\hat{\gamma})\right)\right]^2+\left[\ln\left(\Phi(z_{iT_i}\hat{\gamma})\right)\right]^2\right) \\ \vdots & \vdots & & \vdots \\ -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{iT_i}\hat{\gamma})\right)\right]^2+\left[\ln\left(\Phi(z_{i1}\hat{\gamma})\right)\right]^2\right) & -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{iT_i}\hat{\gamma})\right)\right]^2+\left[\ln\left(\Phi(z_{i2}\hat{\gamma})\right)\right]^2\right) & \cdots & \left\{\ln\left(\Phi(z_{iT_i}\hat{\gamma})\right)\right\}^2\left(1-\frac{2}{T_i}\right) \end{pmatrix}$$

$\hat{\Sigma}_{Ci} = \frac{1}{\left(\widehat{\mu\delta}\right)T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\hat{\gamma})\right)\right\}^2 \iota_i\iota_i'$, where ι_i is a $(T_i \times 1)$ vector of 1’s.

3.6 *Technical Efficiency*

It may be noted that

$$E\left[\frac{1}{\mu}\ln\left(\Phi(z_{it}\gamma)\right)u_{it}\right] = \ln\left(\Phi(z_{it}\gamma)\right) \tag{38}$$

Once the coefficient vector γ is estimated, the mean technical efficiency can be readily estimated as

$$\widehat{TE_{it}} = \exp\left(\ln\left[\Phi(z_{it}\hat{\gamma})\right]\right) = \Phi(z_{it}\hat{\gamma}) \tag{39}$$

4. An Empirical Illustration

Annual data from 1975–1976 to 1984–1985 on farmers from the village of Aurepalle in State of Andhra Pradesh in India[16] are used for empirical illustration. The data are unbalanced for 34 farmers with 271observations over the period of 10 years[17]. This data set was made available to us by Hung-Jen Wang to whom we are thankful. In the past, this dataset has been used in several inefficiency studies including Battese and Coelli (1995), Coelli and Battese (1996) and Wang (2002). In line with these studies, the Cobb-Douglas functional form is chosen for our stochastic production function. For the production function, y_{it}: $\ln(Y_{it})$ where Y is the total value of output (in Rupees, in 1975-76 values) from the crops which are grown; x_{it}: $\{\ln(Land_{it}), \ln(Labor_{it}), \ln(Bullock_{it}), PILand_{it}, \ln[\text{Max}(Cost_{it}, 1 - D_{it})], Year_{it}\}$ where *Land* is the total area of irrigated and unirrigated land operated, *Labor* is the total hours of family and hired labor, *Bullock* is the hours of bullock labor and *PILand* is the proportion of operated land that is irrigated. *Cost* is the value of other inputs, including fertilizer, manure, pesticides, machinery, etc. and *D* is a variable which has a value of one if *Cost* is positive, and a value of zero if otherwise. *Year* is the year of the observation, numbered from 1 to 10, which accounts for the Hicksian neutral technological change. For the efficiency effect specification, z_{it}: $\{Age_{it}, Schooling_{it}, Land_{it,}, Land_{it}^2\}$, where *Age* is the age of the primary decision-maker in the farming operation and *Schooling* is the years of formal schooling of the primary decision maker. We expect the efficiency level of the farms to increase with the level of education of the decision maker. However, it is difficult to predict *a priori* the sign on the effect of age of primary decision maker on efficiency. If the younger people have better knowledge of farming techniques and management then the farms with younger decision makers are likely to be more

[16] These farm-level data on the agricultural operations of farmers were collected by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

[17] This data set contains all the 10 year observations for 16 of the farmers, and 2 minimal observations for 2 of the farmers.

technically efficient, other things remaining the same. On the other hand, if the experience gained over the years matters for farming, then the farms managed by older persons might be technically more efficient. Thus, the effect of age of primary decision maker on technical efficiency is an empirical issue. *Land* and $Land^2$ are used to capture non-linear relationship between efficiency and farm size. There is a very old and vast literature debating the negative relationship between farm size and productivity where the latter is defined as output per land area cultivated (Sen, 1966; Carter, 1984; Eswaran and Kotwal, 1986; Bhalla and Roy, 1988; Benjamin, 1995; Barrett, 1996; Heltberg, 1998). However, the effect of land size on farm efficiency is investigated only recently. Whether small farms have technical efficiency advantage and remain competitive in the light of ongoing transformation of agricultural markets and supply chain, is an empirical question. Using the Mexican panel data on farming, Kagin et al. (2016) find an inverse efficiency relationship with farm size within the stochastic frontier framework of Battese and Coelli (1995). Using the Brazilian farming data, Helfand and Levine (2004) reveal a non-linear relationship between farm size and efficiency, with efficiency first falling and then rising with size. Similarly, for the Swedish dairy farms, Hansson (2008) also reports a U-shaped relationship between efficiency and farm size. The insertion of Land and Land squared terms in the efficiency model allows us to test whether the U-shaped farm size-efficiency relationship also holds for Indian farmers.

The summary statistics of sample data are presented in Table 1. The land area cultivated varies from 0.20 to 20.97 hectares. The percentage of land area under irrigation varies from 0 to 100%. The age of farmers varies from 26 to 90 years and the level of education of farmers varies from illiteracy to 10 years of schooling.

[insert Table 1 here]

The non-linear least squares (NLS) parameter estimates of the proposed model (equation 17) are obtained using *Matlab* software package. These estimates along with their standard errors are presented in cols. 2 and 3 of Table 2. The coefficients of inputs in the production function represent their output elasticities. The output elasticities with respect to *Land* and *Labor* are positive and statistically significant. In terms of the magnitude of elasticity, labor turns out to be most important factor of production. The output elasticity of *Bullock*, which is negative and statistical significant, is not to our expectations. This result was also observed in Battese and Coelli (1992, 1995), Coelli and Battese (1996) and Battese et al. (1989). A plausible explanation for this result, as provided in Battese and Coelli (1995), is that farmers may use bullocks more in years of poor production (associated with low rainfall) for the purpose of weed control, levy bank maintenance etc., which are difficult to conduct in years of higher rainfall and higher output. Hence, the bullock-labor variable may be acting as an inverse proxy for rainfall. The elasticity of *Cost* (of other inputs) is negative but statistically insignificant. The elasticity of *PILand* is positive and significant at the10 percent level, implying that higher the proportion of irrigated farming, the larger is the output, other things remaining the same. The coefficient on Year is positive and significant, implying that there is significant technological progress.

[insert Table 2 here]

In the technical efficiency effects specification, the coefficient of *Schooling* of the decision maker is positive and statistically significant, implying that the efficiency of a farm improves with the level of education of the primary decision maker. The coefficient of *Age* of the decision maker is also positive and significant, implying that, *cetris paribus*, farms managed by older farmers are more efficient than those managed by younger farmers. This is expected because in the traditional farming, the practical experience gained by farmers over the years is likely to improve their farming efficiency. While the coefficient of *Land* is negative (-0.247) and statistically significant at the 1 percent level, the coefficient of *Land Squared is* positive (0.008) and statistically significant at the 10 percent level. This implies that the efficiency relationship with farm size is U-shaped with efficiency first declining with farm size and then increasing with size. This finding is similar to the results reported in Helfand and Levine (2004) for the Brazilian farms. These results suggest that since small farms have efficiency advantage, it could be that a heterogeneous farm structure, in which small farms coexist with large ones, is consistent with promoting agricultural growth. While the small farms' technical efficiency advantage has ramifications for their potential role in combating poverty and enhancing food security, the medium sized farmers should aim for farm sizes which are in the larger farm size segments to take advantage of higher productive efficiency.

The null hypothesis that there are no efficiency effects (i.e., all the coefficients of efficiency effects model are zero) is rejected at the 1% significance level by the Wald statistics. The technical efficiency levels range from 0.344 to 0.989 with an average level of 0.783 (Table 3, col. 2). The estimated probability density function (pdf) of technical efficiency which is skewed to the left, is leptokurtic as revealed by the Kurtosis statistics (Figure 1).

[insert Table 3 here]

[insert Figure 1 here]

The technical efficiency scores of farms (averaged over the sample period) along with their rankings are presented in cols. 2 and 3 of Table 4. It is also worth noting that the average efficiency level of farmers shows a mild increase over time, from an average of 0.777 in the first half of the period to 0.789 in the second half (Table 5).

[insert Table 4 here]

[insert Table 5 here]

The model with efficiency effects specified by a logistic cumulative distribution function (logit model) is also estimated to see the sensitivity of results. The input elasticities of the production function with logit efficiency effects specification presented in col 4 of Table 2 are quite similar to those with the probit specification, in terms of magnitude and signs. The estimated coefficients of *Age* and *Schooling* of the decision maker in the logit specification of efficiency effects have the same signs as observed in the case the probit specification. The efficiency relationship with farm size is also observed to be U-shaped. The average efficiency level of farms based on the logit specification is 0.818 which is slightly higher than that observed in the case of the probit specification (0.783) (Table 3). The efficiency ranking of farms by the logit model is almost the same (with some minor differences) as that by the probit model (Table 4). Like the probit model, the logit specification also shows a mild increase in average efficiency from first half period to the second half (Table 5). It is also worth noting that the correlations between the probit and logit efficiency estimates and their rankings are quite high, 0.998 and 0.997 respectively.

5. **Concluding Remarks**

This paper proposed a stochastic frontier panel data model which accommodates time-invariant unobserved heterogeneity along with efficiency effects. The efficiency effects are specified by a standard normal cumulative distribution function of exogenous variables which ensures the efficiency scores to lie in a unit interval. The model is within-transformed and then estimated with non-linear least squares. The estimated parameters thus obtained are consistent. The JLMS transformation is not required to compute efficiency scores. In our case, the efficiency scores are calculated directly once parameters of the model are obtained.

The empirical exercise conducted with widely used panel data on Indian farmers reveals that both the education and age of the primary decision-maker enhance the efficiency of farms. The relationship between efficiency and farm size is found to be U-shaped. This suggests that since small farms have efficiency advantage, it could be that a heterogeneous farm structure, in which small farms co-exist with large ones, is consistent with promoting agricultural growth.

Appendix A

Proof of Theorem 3:

Using Taylor series expansion we have

$$\left.\frac{\partial Q_L(\theta)}{\partial \theta}\right|_{\hat{\theta}_{NLS}} = \left.\frac{\partial Q_L(\theta)}{\partial \theta}\right|_{\theta_0} + \left.\frac{\partial^2 Q_L(\theta)}{\partial \theta\, \partial \theta'}\right|_{\tilde{\theta}} \left(\hat{\theta}_{NLS} - \theta_0\right) = \mathbf{0} \qquad (A.1)$$

where $\tilde{\theta}$ lies between $\hat{\theta}_{NLS}$ and θ_0. On re-writing (A.1) we obtain

$$\sqrt{L}\left(\hat{\theta}_{NLS} - \theta_0\right) = \left[\frac{1}{L}\left.\frac{\partial^2 Q_L(\theta)}{\partial \theta\, \partial \theta'}\right|_{\tilde{\theta}}\right]^{+} \frac{1}{\sqrt{L}} \left.\frac{\partial Q_L(\theta)}{\partial \theta}\right|_{\theta_0} \qquad (A.2)$$

where + denotes the Moore-Penrose generalized inverse[18].

As $\operatorname{plim} \frac{1}{L}\left.\frac{\partial^2 Q_L(\theta)}{\partial \theta\, \partial \theta'}\right|_{\theta_0} = \operatorname{plim} \frac{1}{L}\left.\frac{\partial \mathbf{g}'}{\partial \theta}\frac{\partial \mathbf{g}}{\partial \theta'}\right|_{\theta_0}$, the conclusion of the theorem follows from assumption (iv) and equations (A.2) and (22) by repeated applications of Theorem 3.2.7 in Amemiya (1985) whose proof can be found in Rao (1973, p. 122).

[18] For any square matrix A there exists a positive constant δ_0 such that for all $\delta > \delta_0$, A+δI is non-singular. Here I is the identity matrix.

Appendix B

Proof of Theorem 4

For ready reference, we re-write the model and the seven assumptions stated in the main text.

$$y_{it} = x_{it}\beta + \ln\left[\Phi(z_{it}\gamma)\right] + \varepsilon_{it} \qquad (B.1)$$

where $\varepsilon_{it} = v_{it} + \frac{1}{\mu}\ln\left[\Phi(z_{it}\gamma)\right]\left(u_{it} - \mu\right)$ is independent but not identically distributed.

Or after within transformation

$$\tilde{y}_{it} = \tilde{x}_{it}\beta + \ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right) + \tilde{\varepsilon}_{it} \quad i = 1,...,N; t = 1,...,T_i \qquad (B.2)$$

where $\bar{w}_i = \frac{1}{T_i}\sum_{p=1}^{T_i} w_{ip}$ and $\tilde{w}_i = w_{it} - \bar{w}_i,\ w \in \{y, x, \varepsilon\}$.

The assumptions are:

(i) $E\left(v_{it}^2\right) = \sigma_v^2$

(ii) $E\left(v_{it}v_{is}\right) = 0\ if\ t \neq s$

(iii) $E\left(v_{it}v_{jt}\right) = 0\ if\ i \neq j$

(iv) $E\left(u_{it}\right) = \mu$

(v) $E\left(u_{it}u_{js}\right) = E\left(u_{it}\right)E\left(u_{js}\right) = \mu^2\ \ \forall\ i \neq j$ and/or $t \neq s$

(vi) $E\left(\left[u_{it} - E\left(u_{it}\right)\right]^2\right) = \sigma_u^2$

(vii) $E\left(v_{it}u_{js}\right) = E\left(v_{it}\right)E\left(u_{js}\right) = 0\ \ \forall\ i, j, t$ and s

In the above setting, consider

$$\begin{aligned}\tilde{\varepsilon}_{it} &= v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik} + \frac{1}{\mu}\ln\left(\Phi(z_{it}\gamma)\right)\left(u_{it}-\mu\right) - \frac{1}{\mu T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\left(u_{ik}-\mu\right)\\
&= v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik} + \frac{1}{\mu}\ln\left(\Phi(z_{it}\gamma)\right)u_{it} - \ln\left(\Phi(z_{it}\gamma)\right) - \left[\frac{1}{\mu T_i}\sum_{k=1}^{T_i}\ln\left([\Phi(z_{ik}\gamma)]\right)u_{ik} - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\right]\\
&= v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik} + \frac{1}{\mu}\left\{\ln\Phi(z_{it}\gamma)u_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\} - \left\{\ln\left(\Phi(z_{it}\gamma)\right) - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\right\}\end{aligned}$$

Taking expectations on both sides we have

$$\begin{aligned}E\left(\tilde{\varepsilon}_{it}\right) &= E\left(v_{it}\right) - \frac{1}{T_i}\sum_{k=1}^{T_i} E\left(v_{ik}\right) + \frac{1}{\mu}\left\{\ln\left([\Phi(z_{it}\gamma)]\right)E\left(u_{it}\right) - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left([\Phi(z_{ik}\gamma)]\right)E\left(u_{ik}\right)\right\}\\
&\quad - \left\{\ln\left([\Phi(z_{it}\gamma)]\right) - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left([\Phi(z_{ik}\gamma)]\right)\right\}\end{aligned}$$

$$= \frac{1}{\mu}\left\{\ln\left([\Phi(z_{it}\gamma)]\right)\mu - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left([\Phi(z_{ik}\gamma)]\right)\mu\right\} - \left\{\ln\left([\Phi(z_{it}\gamma)]\right) - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left([\Phi(z_{ik}\gamma)]\right)\right\} = 0$$

$$E\left(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{js}\right) =$$

$$E\left(\begin{aligned}&\left[v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik} + \frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\} - \left\{\ln\left(\Phi(z_{it}\gamma)\right) - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\right\}\right]\times\\
&\left[v_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j} v_{jp} + \frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{jp}\right\} - \left\{\ln\left(\Phi(z_{js}\gamma)\right) - \frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)\right\}\right]\end{aligned}\right)$$

$$= E\left(\begin{aligned}&\left[v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik} + \frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\} - \left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right]\times\\
&\left[v_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j} v_{jp} + \frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{jp}\right\} - \left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right]\end{aligned}\right)$$

$$= E\left\{\left(v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik}\right)\times\left(v_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j} v_{jp}\right)\right\} \qquad (I)$$

$$+ E\left\{\left\{v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik}\right\}\times\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{jp}\right\}\right]\right\} \qquad (II)$$

$$- E\left\{\left\{v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik}\right\}\times\left[\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right]\right\} \qquad (III)$$

$$+ E\left\{\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times\left\{v_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j} v_{jp}\right\}\right\} \qquad (IV)$$

$$+ E\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times \\ \left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{jp}\right\}\right]\end{array}\right\} \qquad (V)$$

$$- E\left\{\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\} \qquad (VI)$$

$$- E\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left[v_{js} - \frac{1}{T_j}\sum_{p=1}^{T_j} v_{jp}\right]\right\} \qquad (VII)$$

$$-E\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{js}-\frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{jp}\right\}\right]\right\} \quad (VIII)$$

$$+E\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\} \quad (IX)$$

$$(B.3)$$

Given the assumptions (i)-(vii), it is easy to see that in the above equation, the terms (*II*), (*III*), (*IV*) and (*VII*) in (*B*.3) are zero. Hence, we can rewrite (*B.3*) as

$$E\left(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{js}\right)=E\left\{\left(v_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\times\left(v_{js}-\frac{1}{T_j}\sum_{p=1}^{T_j}v_{jp}\right)\right\} \quad (I)$$

$$+E\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times\\\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{js}-\frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{jp}\right\}\right]\end{array}\right\} \quad (V)$$

$$-E\left\{\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\} \quad (VI)$$

$$-E\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{js}-\frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{jp}\right\}\right]\right\} \quad (VIII)$$

$$+\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\} \quad (IX)$$

$$(B.4)$$

Taking expectations on both sides in (*VI*) we get

$$\left\{\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)E\left(u_{it}\right)-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)E\left(u_{ik}\right)\right\}\right]\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}$$

$$=\left\{\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)\mu-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\mu\right\}\right]\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}$$

$$=\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\} \quad (B.5)$$

Similarly, taking expectations on both sides in (*VIII*) we get

$$\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)E\left(u_{js}\right)-\frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)E\left(u_{jp}\right)\right\}\right]\right\}$$

$$=\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\} \qquad (B.6)$$

In the light of (*B.5*) and (*B.6*), (*B.4*) can be re-written as

$$E\left(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{js}\right)=E\left\{\left(v_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\times\left(v_{js}-\frac{1}{T_j}\sum_{p=1}^{T_j}v_{jp}\right)\right\}$$

$$+E\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times\\\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{js}-\frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{jp}\right\}\right]\end{array}\right\}$$

$$-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\} \qquad (B.7)$$

In (*B.7*) four cases arise.

- (i) $t=s, i=j$
- (ii) $t\neq s, i=j$
- (iii) $i\neq j, t=s$
- (iv) $i\neq j, t\neq s$

Case (i) corresponds to the same farm in same time period, whereas case (ii) corresponds to the same farm in different time periods. Case (iii) corresponds to different farms in the same period of time, and case (iv) corresponds to different farms across different time period.

Case (i): Given (*B.7*), we have

$$E\left(\tilde{\varepsilon}_{it}^2\right)=E\left\{\left(v_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)^2\right\}$$

$$+E\left\{\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]^2\right\}$$

$$-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\}$$

$$=E\left\{v_{it}^2+\left(\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)^2-2v_{it}\left(\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\right\}$$

$$+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]E\left\{u_{it}^2\right\}-\frac{2}{\mu^2T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{ik}\gamma)\right)\right]E\left\{u_{it}u_{ik}\right\}$$

$$+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\sum_{p=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)E\left\{u_{ik}u_{ip}\right\}-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2$$

$$=\left\{E\left(v_{it}^2\right)+E\left(\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)^2-E\left(2v_{it}\times\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\right\}$$

$$+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]E\left\{u_{it}^2\right\}-\frac{2}{\mu^2T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]E\left\{u_{it}^2\right\}$$

$$-\frac{2}{\mu^2T_i}\sum_{k\neq t}\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{ik}\gamma)\right)\right]E\left\{u_{it}u_{ik}\right\}$$

$$+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2E\left\{u_{ik}^2\right\}+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\sum_{p=1,k\neq p}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)E\left\{u_{ik}u_{ip}\right\}$$

$$-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2$$

$$=\left\{\sigma_v^2+E\left(\frac{1}{T_i^2}\sum_{k=1}^{T_i}v_{ik}^2\right)-E\left(2v_{it}\times\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\right\}$$

$$+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]E\left\{u_{it}^2\right\}-\frac{2}{\mu^2T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]E\left\{u_{it}^2\right\}$$

$$-\frac{2}{\mu^2T_i}\sum_{k\neq t}\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{ik}\gamma)\right)\right]\mu^2$$

$$+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2E\left\{u_{ik}^2\right\}+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\sum_{p=1,k\neq p}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)\mu^2$$

$$-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2$$

We know that $E\left\{u_{it}^2\right\}=\sigma_u^2+\mu^2$. So we have

$$E\left(\tilde{\varepsilon}_{it}^2\right) = \sigma_v^2 + \frac{1}{T_i^2}\sum_{k=1}^{T_i} E\left(v_{ik}^2\right) - \frac{2}{T_i} E\left(v_{it} \times \sum_{k=1}^{T_i} v_{ik}\right)$$
$$+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\left(\sigma_u^2 + \mu^2\right) - \frac{2}{\mu^2 T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\left(\sigma_u^2 + \mu^2\right)$$
$$-\frac{2}{T_i}\sum_{k \neq t}\left[\ln\left(\Phi(z_{it}\gamma)\right) \times \ln\left(\Phi(z_{ik}\gamma)\right)\right]$$
$$+\frac{1}{\mu^2 T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2\left(\sigma_u^2 + \mu^2\right) + \frac{1}{T_i^2}\sum_{k=1}^{T_i}\sum_{p=1,k\neq p}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)$$
$$-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2$$

$$= \sigma_v^2 + \frac{1}{T_i^2}\sum_{k=1}^{T_i}\sigma_v^2 - \frac{2}{T_i}E\left(v_{it}^2\right)$$
$$+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2 + \left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right] - \frac{2}{\mu^2 T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2 - \frac{2}{T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]$$
$$-\frac{2}{T_i}\sum_{k \neq t}\left[\ln\left(\Phi(z_{it}\gamma)\right) \times \ln\left(\Phi(z_{ik}\gamma)\right)\right]$$
$$+\frac{1}{\mu^2 T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2\sigma_u^2 + \frac{1}{T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2 + \frac{1}{T_i^2}\sum_{k=1}^{T_i}\sum_{p=1,k\neq p}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)$$
$$-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2$$

$$
=\frac{T_i-1}{T_i}\sigma_v^2-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2
$$

$$
+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2-\frac{2}{\mu^2T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2\sigma_u^2
$$

$$
+\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]-\frac{1}{T_i}\ln\left(\Phi(z_{it}\gamma)\right)\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]
$$

$$
-\frac{1}{T_i}\ln\left(\Phi(z_{it}\gamma)\right)\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]+\frac{1}{T_i^2}\sum_{k=1}^{T_i}\sum_{p=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)
$$

$$
=\frac{T_i-1}{T_i}\sigma_v^2-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2
$$

$$
+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2-\frac{2}{\mu^2T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2\sigma_u^2
$$

$$
+\ln\left(\Phi(z_{it}\gamma)\right)\left\{\ln\left(\Phi(z_{it}\gamma)\right)-\frac{1}{T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]\right\}
$$

$$
-\frac{1}{T_i}\ln\left(\Phi(z_{it}\gamma)\right)\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]+\frac{1}{T_i^2}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]\times\sum_{p=1}^{T_i}\left[\ln\left(\Phi(z_{ip}\gamma)\right)\right]
$$

$$
=\frac{T_i-1}{T_i}\sigma_v^2-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2
$$

$$
+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2-\frac{2}{\mu^2T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2\sigma_u^2
$$

$$
+\ln\left(\Phi(z_{it}\gamma)\right)\left\{\ln\left(\Phi(z_{it}\gamma)\right)-\frac{1}{T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]\right\}
$$

$$
-\frac{1}{T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]\left\{\ln\left(\Phi(z_{it}\gamma)\right)-\frac{1}{T_i}\sum_{p=1}^{T_i}\left[\ln\left(\Phi(z_{ip}\gamma)\right)\right]\right\}
$$

$$=\frac{T_i-1}{T_i}\sigma_v^2-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2$$

$$+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2-\frac{2}{\mu^2T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2\sigma_u^2$$

$$+\ln\left(\Phi(z_{it}\gamma)\right)\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}-\frac{1}{T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}$$

$$=\frac{T_i-1}{T_i}\sigma_v^2-\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2+\frac{1}{\mu^2}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2-\frac{2}{\mu^2T_i}\left[\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\right]\sigma_u^2$$

$$+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2\sigma_u^2+\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}^2$$

$$=\frac{T_i-1}{T_i}\sigma_v^2+\frac{\sigma_u^2}{\mu^2}\left\{\left\{\ln\left(\Phi(z_{it}\gamma)\right)\right\}^2\left(1-\frac{2}{T_i}\right)+\frac{1}{T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2\right\} \qquad (B.8)$$

Case (ii): Again given (*B.7*), we have

$$E\left(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{is}\right)=$$

$$E\left\{\left(v_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\times\left(v_{is}-\frac{1}{T_i}\sum_{p=1}^{T_j}v_{ip}\right)\right\}-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\}$$

$$+E\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times\\\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{is}\gamma)\right)u_{is}-\frac{1}{T_j}\sum_{p=1}^{T_i}\ln\left(\Phi(z_{ip}\gamma)\right)u_{ip}\right\}\right]\end{array}\right\}$$

$$=E\left(v_{it}v_{is}\right)-E\left(v_{it}\frac{1}{T_i}\sum_{p=1}^{T_i}v_{ip}\right)-E\left(v_{is}\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)+E\left(\left(\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\left(\frac{1}{T_i}\sum_{p=1}^{T_i}v_{ip}\right)\right)$$

$$-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\}$$

$$+\frac{1}{\mu^2}\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{is}\gamma)\right)\right]E\left\{u_{it}u_{is}\right\}-\frac{1}{\mu^2T_i}\sum_{p=1}^{T_i}\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{ip}\gamma)\right)\right]E\left\{u_{it}u_{ip}\right\}$$

$$-\frac{1}{\mu^2T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\times\ln\left(\Phi(z_{is}\gamma)\right)\right]E\left\{u_{ik}u_{is}\right\}+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\sum_{p=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)E\left\{u_{ik}u_{ip}\right\}$$

$$=0-\frac{1}{T_i}E\left(v_{it}^2\right)-\frac{1}{T_i}E\left(v_{is}^2\right)+\frac{1}{T_i^2}E\left(\left(\sum_{k=1}^{T_i}\left(v_{ik}^2\right)\right)\right)-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\}$$

$$+\frac{1}{\mu^2}\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{is}\gamma)\right)\right]\mu^2-\frac{1}{\mu^2T_i}\sum_{p\neq t}\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{ip}\gamma)\right)\right]\mu^2$$

$$-\frac{1}{\mu^2T_i}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^2E\left\{u_{it}^2\right\}-\frac{1}{\mu^2T_i}\sum_{k\neq s}\left[\ln\left(\Phi(z_{is}\gamma)\right)\times\ln\left(\Phi(z_{ik}\gamma)\right)\right]\mu^2$$

$$-\frac{1}{\mu^2T_i}\left[\ln\left(\Phi(z_{is}\gamma)\right)\right]^2E\left\{u_{is}^2\right\}+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2E\left\{u_{ik}^2\right\}$$

$$+\frac{1}{\mu^2T_i^2}\sum_{k=1}^{T_i}\sum_{p=1,k\neq p}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)\mu^2$$

Since $E\left\{u_{it}^2\right\}=\sigma_u^2+\mu^2$, we have

$$=-\frac{2}{T_i}\sigma_v^2+\frac{1}{T_i^2}\sum_{k=1}^{T_i}\sigma_v^2-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\}$$

$$+\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{is}\gamma)\right)\right]-\frac{1}{T_i}\sum_{p=1}^{T}\left[\ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\Phi(z_{ip}\gamma)\right)\right]-\frac{\sigma_u^2}{\mu^2T_i}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^2$$

$$-\frac{1}{T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{is}\gamma)\right)\times\ln\left(\Phi(z_{ik}\gamma)\right)\right]-\frac{\sigma_u^2}{\mu^2T_i}\left[\ln\left(\Phi(z_{is}\gamma)\right)\right]^2+\frac{\sigma_u^2}{\mu^2T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2$$

$$+\frac{1}{T_i^2}\sum_{k=1}^{T_i}\sum_{p=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right)$$

$$
\begin{aligned}
&= -\frac{1}{T_i}\sigma_v^2 - \left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\} + \ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right) \\
&-\frac{1}{T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{is}\gamma)\right)\times\ln\left(\Phi(z_{ik}\gamma)\right)\right] + \frac{1}{T_i^2}\sum_{k=1}^{T_i}\sum_{p=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\ln\left(\Phi(z_{ip}\gamma)\right) \\
&-\frac{\sigma_u^2}{\mu^2 T_i}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^2 - \frac{\sigma_u^2}{\mu^2 T_i}\left[\ln\left(\Phi(z_{is}\gamma)\right)\right]^2 + \frac{\sigma_u^2}{\mu^2 T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{1}{T_i}\sigma_v^2 - \left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\} + \ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right) \\
&-\frac{1}{T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]\left[\ln\left(\Phi(z_{is}\gamma)\right) - \frac{1}{T_i}\sum_{p=1}^{T_i}\left[\ln\left(\Phi(z_{ip}\gamma)\right)\right]\right] - \frac{\sigma_u^2}{\mu^2 T_i}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^2 \\
&-\frac{\sigma_u^2}{\mu^2 T_i}\left[\ln\left(\Phi(z_{is}\gamma)\right)\right]^2 + \frac{\sigma_u^2}{\mu^2 T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{1}{T_i}\sigma_v^2 - \left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\right\} + \ln\left(\Phi(z_{it}\gamma)\right)\times\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right) \\
&-\frac{1}{T_i}\sum_{k=1}^{T_i}\left[\ln\left(\Phi(z_{ik}\gamma)\right)\right]\times\ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i}\Phi(z_{ip}\gamma)\right]^{\frac{1}{T_i}}}\right) - \frac{\sigma_u^2}{\mu^2 T_i}\left[\ln\left(\Phi(z_{it}\gamma)\right)\right]^2 - \frac{\sigma_u^2}{\mu^2 T_i}\left[\ln\left(\Phi(z_{is}\gamma)\right)\right]^2 \\
&+\frac{\sigma_u^2}{\mu^2 T_i^2}\sum_{k=1}^{T_i}\left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2
\end{aligned}
$$

$$= -\frac{1}{T_i}\sigma_v^2 - \left\{ \left\{ \ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i} \Phi(z_{ik}\gamma) \right]^{\frac{1}{T_i}}} \right) \right\} \times \left\{ \ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i} \Phi(z_{ip}\gamma) \right]^{\frac{1}{T_i}}} \right) \right\} \right\}$$

$$+ \left\{ \left\{ \ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i} \Phi(z_{ik}\gamma) \right]^{\frac{1}{T_i}}} \right) \right\} \times \left\{ \ln\left(\frac{\Phi(z_{is}\gamma)}{\left[\prod_{p=1}^{T_i} \Phi(z_{ip}\gamma) \right]^{\frac{1}{T_i}}} \right) \right\} \right\}$$

$$- \frac{\sigma_u^2}{\mu^2 T_i}\left[\ln\left(\Phi(z_{it}\gamma) \right) \right]^2 - \frac{\sigma_u^2}{\mu^2 T_i}\left[\ln\left(\Phi(z_{is}\gamma) \right) \right]^2 + \frac{\sigma_u^2}{\mu^2 T_i^2}\sum_{k=1}^{T_i}\left\{ \ln\left(\Phi(z_{ik}\gamma) \right) \right\}^2$$

We then have

$$E\left(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{is} \right) = -\frac{1}{T_i}\sigma_v^2 + \frac{\sigma_u^2}{\mu^2}\left\{ \frac{1}{T_i^2}\sum_{k=1}^{T_i}\left\{ \ln\left(\Phi(z_{ik}\gamma) \right) \right\}^2 - \frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{it}\gamma) \right) \right]^2 + \left[\ln\left(\Phi(z_{is}\gamma) \right) \right]^2 \right) \right\} \qquad (B.9)$$

Case (iii): Again from (*B.7*), we have

$$E\left(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{jt}\right) = E\left(\left(v_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik}\right)\times\left(v_{jt} - \frac{1}{T_j}\sum_{p=1}^{T_j} v_{jp}\right)\right)$$

$$-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{jt}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}$$

$$+E\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it} - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times \\ \left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{jt}\gamma)\right)u_{jt} - \frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{ip}\right\}\right]\end{array}\right\}$$

$$= E\left(v_{it}v_{jt}\right) - E\left(v_{it}\frac{1}{T_j}\sum_{p=1}^{T_j} v_{jp}\right) - E\left(v_{jt}\frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik}\right) + E\left(\left(\frac{1}{T_i}\sum_{k=1}^{T_i} v_{ik}\right)\left(\frac{1}{T_j}\sum_{p=1}^{T_j} v_{jp}\right)\right)$$

$$-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{jt}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}$$

$$+\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)E\left(u_{it}\right) - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)E\left(u_{ik}\right)\right\}\right]\times \\ \left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{jt}\gamma)\right)E\left(u_{jt}\right) - \frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)E\left(u_{ip}\right)\right\}\right]\end{array}\right\}$$

$$
\begin{aligned}
&= E\left(v_{it}v_{jt}\right)-\frac{1}{T_j}\sum_{p=1}^{T_j}E\left(v_{it}v_{jp}\right)-\frac{1}{T_i}\sum_{k=1}^{T_i}E\left(v_{jt}v_{ik}\right)+\frac{1}{T_iT_j}\left(\sum_{k=1}^{T_i}\sum_{p=1}^{T_j}E\left(v_{ik}v_{jp}\right)\right)\\
&-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{jt}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}\\
&+\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)\mu-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\mu\right\}\right]\times\\ \left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{jt}\gamma)\right)\mu-\frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)\mu\right\}\right]\end{array}\right\}\\
&= 0-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{jt}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}\\
&+\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{jt}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}=0 \qquad (B.10)
\end{aligned}
$$

Case (iv): Re-writing (*B.7*) we have

$$E\left(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{js}\right)=E\left(\left(v_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\times\left(v_{js}-\frac{1}{T_j}\sum_{p=1}^{T_j}v_{jp}\right)\right)$$

$$-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}$$

$$+E\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)u_{it}-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)u_{ik}\right\}\right]\times\\ \left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)u_{jt}-\frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)u_{ip}\right\}\right]\end{array}\right\}$$

$$=E\left(v_{it}v_{js}\right)-E\left(v_{it}\frac{1}{T_j}\sum_{p=1}^{T_j}v_{jp}\right)-E\left(v_{js}\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)+E\left(\left(\frac{1}{T_i}\sum_{k=1}^{T_i}v_{ik}\right)\left(\frac{1}{T_j}\sum_{p=1}^{T_j}v_{jp}\right)\right)$$

$$-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\}\times\left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}$$

$$+\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)E\left(u_{it}\right)-\frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)E\left(u_{ik}\right)\right\}\right]\times\\ \left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)E\left(u_{js}\right)-\frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)E\left(u_{jp}\right)\right\}\right]\end{array}\right\}$$

$$= E\left(v_{it}v_{js}\right) - \frac{1}{T_j}\sum_{p=1}^{T_j} E\left(v_{it}v_{jp}\right) - \frac{1}{T_i}\sum_{k=1}^{T_i} E\left(v_{js}v_{ik}\right) + \frac{1}{T_i T_j}\left(\sum_{k=1}^{T_i}\sum_{p=1}^{T_j} E\left(v_{ik}v_{jp}\right)\right)$$

$$-\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\} \times \left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}$$

$$+\left\{\begin{array}{l}\left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{it}\gamma)\right)\mu - \frac{1}{T_i}\sum_{k=1}^{T_i}\ln\left(\Phi(z_{ik}\gamma)\right)\mu\right\}\right] \times \\ \left[\frac{1}{\mu}\left\{\ln\left(\Phi(z_{js}\gamma)\right)\mu - \frac{1}{T_j}\sum_{p=1}^{T_j}\ln\left(\Phi(z_{jp}\gamma)\right)\mu\right\}\right]\end{array}\right\}$$

$$= 0 - \left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\} \times \left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\}$$

$$+\left\{\left\{\ln\left(\frac{\Phi(z_{it}\gamma)}{\left[\prod_{k=1}^{T_i}\Phi(z_{ik}\gamma)\right]^{\frac{1}{T_i}}}\right)\right\} \times \left\{\ln\left(\frac{\Phi(z_{js}\gamma)}{\left[\prod_{p=1}^{T_j}\Phi(z_{jp}\gamma)\right]^{\frac{1}{T_j}}}\right)\right\}\right\} = 0 \qquad (B.11)$$

If we stack firm observations together then we have the following structure of covariance matrix.

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 & \cdots & 0 & 0 \\ 0 & \vdots & \vdots & 0 & 0 \\ 0 & 0 & \Sigma_i & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \Sigma_N \end{pmatrix} \qquad (B.12)$$

which has a dimension of $\left(\sum_{i=1}^{N} T_i \times \sum_{i=1}^{N} T_{i_i}\right)$.

$$\Sigma_i = \Sigma_{Ai} + \Sigma_{Bi} + \Sigma_{Ci} \text{ has a dimension of } \left(T_i \times T_i\right). \qquad (B.13)$$

In the above equation,

$$\Sigma_{Ai} = \sigma_v^2 \begin{pmatrix} \frac{T_i - 1}{T_i} & \frac{-1}{T_i} & \cdots & \frac{-1}{T_i} \\ \frac{-1}{T_i} & \frac{T_i - 1}{T_i} & \cdots & \frac{-1}{T_i} \\ \vdots & \vdots & & \vdots \\ \frac{-1}{T_i} & & \cdots & \frac{T_i - 1}{T_i} \end{pmatrix},$$

$$\Sigma_{Bi} =$$

$$\frac{\sigma_u^2}{\mu^2} \begin{pmatrix} \left\{\ln\left(\Phi(z_{i1}\gamma)\right)\right\}^2 \left(1 - \frac{2}{T_i}\right) & -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{i1}\gamma)\right)\right]^2 + \left[\ln\left(\Phi(z_{i2}\gamma)\right)\right]^2\right) & \cdots & -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{i1}\gamma)\right)\right]^2 + \left[\ln\left(\Phi(z_{iT_i}\gamma)\right)\right]^2\right) \\ -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{i2}\gamma)\right)\right]^2 + \left[\ln\left(\Phi(z_{i1}\gamma)\right)\right]^2\right) & \left\{\ln\left(\Phi(z_{i2}\gamma)\right)\right\}^2 \left(1 - \frac{2}{T_i}\right) & \cdots & -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{i2}\gamma)\right)\right]^2 + \left[\ln\left(\Phi(z_{iT_i}\gamma)\right)\right]^2\right) \\ \vdots & \vdots & & \vdots \\ -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{iT_i}\gamma)\right)\right]^2 + \left[\ln\left(\Phi(z_{i1}\gamma)\right)\right]^2\right) & -\frac{1}{T_i}\left(\left[\ln\left(\Phi(z_{iT_i}\gamma)\right)\right]^2 + \left[\ln\left(\Phi(z_{i2}\gamma)\right)\right]^2\right) & \cdots & \left\{\ln\left(\Phi(z_{iT_i}\gamma)\right)\right\}^2 \left(1 - \frac{2}{T_i}\right) \end{pmatrix}$$

$\Sigma_{Ci} = \frac{\sigma_u^2}{\mu^2 T_i^2} \sum_{k=1}^{T_i} \left\{\ln\left(\Phi(z_{ik}\gamma)\right)\right\}^2 \iota_i \iota_i'$, where ι_i is a $\left(T_i \times 1\right)$ vector of 1's.

References

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

Alvarez, A., Amsler, C., Orea, L., Schmidt, P., 2006. Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics. *Journal of Productivity Analysis*. 25, 201–212.

Amemiya, T., 1985. Advanced Econometrics. Cambridge, MA: Harvard University Press.

Barrett, C.B., 1996. On price risk and the inverse farm size–productivity relationship. *Journal of Development Economics*. 51,193-215.

Battese, G., Coelli, T., Colby, T., 1989. Estimation of frontier production functions and the efficiencies of Indian farms using panel data from ICRISTAT's village level studies. *Journal of Quantitative Economics*. 5, 327-348.

Battese, G.E., Coelli, T.J., 1988. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, 30, 387–399.

Battese, G., Coelli, T., 1992. Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis*. 3, 153-169.

Battese, G.E., Coelli, T.J., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*. 20, 325-332.

Belotti, F., Ilardi, G., 2018. Consistent inference in fixed-effects stochastic frontier models. *Journal of Econometrics*. 202, 161-177.

Benjamin, D., 1995. Can unobserved land quality explain the inverse productivity relationship? *Journal of Development Economics*. 46, 51–84.

Bhalla, S. S., Roy, P., 1988. Mis-specification in farm productivity analysis: The role of land quality. *Oxford Economic Papers*. 40, 55-73.

Carter, M., 1984. Identification of the inverse relationship between farm size and productivity: An empirical analysis of peasant agricultural production. *Oxford Economic Papers*. 36, 131–146.

Caudill, S.B., Ford, J.M., 1993. Biases in frontier estimation due to heteroscedasticity. *Economic Letters*. 41, 17–20.

Caudill, S.B., Ford, J.M., Gropper, D.M., 1995. Frontier estimation and firm-specific inefficiency measures in the presence of heteroscedasticity. *Journal of Business & Economic Statistics*. 13, 105–111.

Chen, Y., Schmidt, P., Wang, H., 2014. Consistent estimation of the fixed effects stochastic frontier model. *Journal of Econometrics*. 181, 65-76.

Coelli, T.J., Battese, G.E., 1996. Identification of factors which influence the technical inefficiency of Indian farmers. *Australian Journal of Agricultural Economics*. 40, 103-28.

Colombi, R., Kumbhakar, S.C., Martini, G., Vittadini, G., 2014. Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. *Journal of Productivity Analysis*. 42, 123-136.

Deprins, D., 1989. Estimation de frontieres de production et Mesures de l'Efficacite Technique. Louvain-la-Neuve, Belgium: CIACO.

Deprins, P., Simar, L., 1989a. Estimating technical efficiencies with corrections for environmental conditions with an application to railway companies. *Annals of Public and Cooperative Economics*. 60, 81-102.

Deprins, P., Simar, L., 1989b. Estimation de frontieres deterministes avec factuers exogenes d'inefficacite. *Annales d'Economie et de Statistique*. 14, 117–150.

Eswaran, M., Kotwal, A., 1986. Access to capital and agrarian production organization. *Economic Journal*. 96, 482-498.

Filippini, M., Greene, W., 2016. Persistent and transient productive inefficiency a maximum simulated likelihood approach. *Journal of Productivity Analysis*. 45, 187–196.

Gonzalez-Farias, G., Dominguez-Molina, J., Gupta, A., 2004. Additive properties of skew normal random vectors. *Journal of Statistical Planning and Inference*. 126, 521-534.

Greene, W.H., 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics.* 126, 269-303.

Hadri, K., 1999. Estimation of a doubly heteroskedastic stochastic frontier cost function. *Journal of Business & Economic Statistics.* 17, 359–363.

Hansson, H., 2008. Are larger farms more efficient? A farm level study of the relationships between efficiency and size on specialized dairy farms in Sweden. *Agricultural and Food Science*. 17, 325-337.

Helfand, S.M., Levine, E.S., 2004. Farm size and the determinants of productive efficiency in the Brazilian Center-West. *Agricultural Economics*. 31, 241-249.

Heltberg, R., 1998. Rural market imperfections and the farm-size productivity relationship: Evidence from Pakistan. *World Development*. 26, 1807-1826.

Horrace, W., Schmidt, P., 1996. Confidence statements for efficiency estimates from stochastic frontier models. *Journal of Productivity Analysis*, 7, 257-282.

Huang, C.J., Liu, J.T., 1994. Estimation of a non-neutral stochastic frontier production function. *Journal of Productivity Analysis*. 5, 171–180.

Jondrow, J., Lovell, C.A.K., Materov, I.S., Schmidt P., 1982. On the estimation of technical inefficiency in stochastic frontier production function model. *Journal of Econometrics*. 19, 233–238.

Kagin, J., Taylor, J.E., Yúnez-Naude, A., 2016. Inverse productivity or inverse efficiency? Evidence from Mexico. *Journal of Development Studies*. 52, 396-411.

Kalirajan, K., 1981. An econometric analysis of yield variability in paddy production. *Canadian Journal of Agricultural Economics*. 29, 283-294.

Kumbhakar, S.C., 1990. Production frontiers, panel data, and time-varying technical inefficiency, *Journal of Econometrics*. 46, 201-212.

Kumbhakar, S.C., 1991. Estimation of technical inefficiency in panel data models with firm and time-specific effects. *Economics Letters*. 36, 43-48.

Kumbhakar, S.C., Ghosh, S., McGuckin, J.T., 1991. A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms. *Journal of Business & Economic Statistics*. 9, 279-286.

Kumbhakar, S.C., Heshmati, A., 1995. Efficiency measurement in Swedish dairy farms: An application of rotating panel data, 1976-88. *American Journal of Agricultural Economics*. 77, 660-674.

Kumbhakar, S.C., Hjalmarsson L., 1993. Technical efficiency and technical progress in Swedish dairy farms. In: Fried HO, Lovell CAK, Schmidt SS (Eds) *The Measurement of Productive Efficiency-Techniques and Applications*, Oxford University Press. 256-270

Kumbhakar, S.C., Hjalmarsson, L., 1995. Labour-use efficiency in Swedish social insurance offices. Journal of Applied Econometrics. 10, 33-47.

Kumbhakar, S., Lovell, C.A.K., 2000. *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge, UK.

Kumbhakar, S. C., Lien G., Hardaker J. B., 2012. Technical efficiency in competing panel data models: a study of Norwegian grain farming. *Journal of Productivity Analysis*. 41, 321-337.

Lancaster, T., 2000. The incidental parameters problem since 1948. *Journal of Econometrics*. 95, 391-414.

Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*. 18, 435-444.

Newey, W., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R.F., McFadden, D. (Eds.), *Handbook of Econometrics*, 4, 2111-2245.

Neyman, J., Scott, E., 1948. Consistent estimates based on partially consistent observations. *Econometrica*. 16, 1-32.

Parmeter, C.F., Kumbhakar, S.C., 2014. Efficiency analysis: A primer on recent advances. *Foundations and Trends(R) in Econometrics,* Now publishers, 7, 191-385.

Parmeter, C.F., Wang, H.J., Kumbhakar, S.C., 2017. Nonparametric estimation of the determinants of inefficiency. *Journal of Productivity Analysis*, 47, 205-221.

Paul, S., Shankar, S., 2018. On estimating efficiency effects in a stochastic frontier model. *European Journal of Operational Research*, 271, 769-774.

Pitt, M.M., Lee, M.F., 1981. The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics*. 9, 43-64.

Rao, C.R., 1973. Linear Statistical Inference and its Applications, Second edition, John Wiley & Sons, New York.

Robinson, P.M., 1988. Root-N-consistent semiparametric regression. *Econometrica*. 56, 931-954.

Schmidt, P., Sickles, R.C., 1984. Production frontiers and panel data. *Journal of Business & Economic Statistics*. 2, 367–374.

Sen, A., 1966. Peasants and dualism with or without surplus labor. *The Journal of Political Economy*. 74, 425-450.

Simar, L., Lovell, C.A.K., van den Eeckaut, P., 1994. Stochastic frontiers incorporating exogenous influences on efficiency. *Discussion Papers No. 9403*, Institut de Statistique, Universite de Louvain.

Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, production processes. *Journal of Econometrics*. 136, 31-64.

Tsionas, E.G., Kumbhakar, S.C., 2012. Firm heterogeneity, persistent and transient technical inefficiency: a generalized true random-effects model. *Journal of Applied Econometrics*. 29, 110-132.

Wang, H.J., 2002. Heteroscedasticity and non-monotonic efficiency effects of a stochastic frontier model. *Journal of Productivity Analysis*. 18, 241-253.

Wang, H.J., Ho, C.W., 2010. Estimating fixed-effect panel stochastic frontier models by model transformation. *Journal of Econometrics*. 157, 286-296.

Wang, H.J., Schmidt, P., 2002. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*. 18, 129-144.

Wikstrom, D., 2015. Consistent method of moments estimation of the true fixed effects model. *Economics Letters*. 137, 62–69.

Table 1: Summary Statistics of Data

	Mean	Maximum	Minimum	Std. Dev.	Observations
Y: Value of output (Rupees)	3705.74	18094.19	36.1133	4565.74	271
Land (hectares)	4.31	20.97	0.20	3.87	271
Labor (hours)	2217.97	12916.00	26.00	2750.50	271
Bullock (hours)	530.97	4316.00	8.00	606.00	271
Cost of other inputs (Rupees)	655.23	6204.99	0	983.44	271
Age of farmer (years)	53.88	90.00	26.0	12.57	271
Schooling of farmer (years)	2.02	10.00	0	2.88	271
PILand	0.14	1.00	0	0.21	271

Table 2: Estimated Stochastic Frontiers and Technical Efficiency Effects

	Model with Probit Efficiency Effects		Model with Logit Efficiency Effects	
Variable	Coefficient	Std. Error[a]	Coefficient	Std. Error[a]
(1)	(2)	(3)	(4)	(5)
Frontier Function				
ln(Land)	0.457***	0.055	0.468***	0.030
ln(Labor)	1.145***	0.063	1.145***	0.042
ln(Bullock)	-0.495***	0.051	-0.495***	0.050
ln(Cost)	-0.002	0.011	-0.002	0.010
PILand	0.264*	0.139	0.260**	0.116
t	0.036***	0.007	0.035***	0.007
	Efficiency Effects		Efficiency Effects	
Constant (γ_0)	0.730	0.825	0.819***	0.274
Age (γ_1)	0.015**	0.002	0.023***	0.007
Schooling (γ_2)	0.125***	0.013	0.187***	0.043
Land (γ_3)	-0.274***	0.094	-0.401***	0.093
Land^2 (γ_4)	0.008*	0.005	0.012***	0.004
Wald statistics[a]:	269.9***		61.2***	
$\hat{\sigma}_v^2$	0.088		0.083	
Observations	271		271	

[a] The Wald statistics has approximately chi-square distribution with degrees of freedom equal to the number of parameters assumed to be zero in the null hypothesis, *H_0*. In the probit and logit models $H_0 : \gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$.

***, ** and * represents significance level at 1, 5 and 10 percent respectively.

Table 3: Summary Statistics of Estimated Technical Efficiency

	Probit Specification	Logit Specification
Mean	0.783	0.818
Median	0.815	0.871
Maximum	0.989	0.999
Minimum	0.344	0.220
Std. Dev.	0.131	0.165
Skewness	-1.027	-1.401
Kurtosis	3.798	4.745
Observations	271	271

Table 4: Farm-Wise Estimates of Mean Technical Efficiency

Farm code	Probit Specification		Logit Specification	
	Estimate	Ranking	Estimate	Ranking
(1)	(2)	(3)	(4)	(5)
1	0.899	7	0.955	7
2	0.858	12	0.917	10
3	0.918	4	0.969	4
4	0.899	6	0.956	5
5	0.691	28	0.702	29
6	0.766	25	0.813	25
7	0.853	14	0.912	14
8	0.846	16	0.898	16
9	0.890	9	0.946	9
10	0.846	15	0.904	15
11	0.820	21	0.867	21
12	0.858	11	0.916	12
13	0.901	5	0.955	6
14	0.839	18	0.896	17
15	0.926	3	0.976	3
16	0.855	13	0.917	11
17	0.960	2	0.991	2
18	0.893	8	0.952	8
19	0.837	19	0.892	18
20	0.522	34	0.467	34
21	0.788	24	0.842	24
22	0.726	26	0.755	26
23	0.810	22	0.867	22
24	0.589	32	0.564	33
25	0.841	17	0.892	19
26	0.688	29	0.710	28
27	0.820	20	0.879	20
28	0.810	23	0.857	23

29	0.624	31	0.617	31
30	0.670	30	0.685	30
31	0.584	33	0.564	32
32	0.704	27	0.726	27
33	0.870	10	0.914	13
34	0.968	1	0.994	1

Table 5: Year-Wise Mean Technical Efficiency Levels

Year (1)	Probit (2)	Logit (3)
1	0.759	0.788
2	0.722	0.739
3	0.806	0.850
4	0.789	0.825
5	0.811	0.846
6	0.783	0.818
7	0.779	0.811
8	0.787	0.825
9	0.779	0.819
10	0.816	0.863
1-5 Years	0.777	0.810
6-10 Years	0.789	0.827

Figure 1: Distribution of Technical Efficiency Scores