



ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΠΑΤΡΩΝ  
UNIVERSITY OF PATRAS

*LECTURE 3- PRODUCTION, TECHNOLOGY  
AND COST FUNCTIONS (TRANSLOG  
PRODUCTION FUNCTION AND ESTIMATION  
OF EFFICIENCY)*

---

**Konstantinos Kounetas**

**School of Business Administration**

**Department of Economics**

**Master of Science in Applied Economic Analysis**

# Introduction

- Please recall in your mind the basic definitions of technical, allocative and total efficiency.
- In the relevant literature, there are two distinct approaches to estimate technical and scale efficiency, stochastic frontier analysis (SFA) and data envelopment analysis (DEA). The main difference between these two methodological approaches is that the stochastic frontier (SF), given that it is parametric, allows the coexistence of inefficiencies and random errors, while the DEA, being nonparametric, attributes the total deviation from the frontier to inefficiency.
- The approach explicitly recognizes that production function represents technically maximum feasible output level for a given level of output.
- The Stochastic Frontier Analysis (SFA) technique may be used in modelling functional relationships where you have theoretical bounds:
  - Estimation of cost functions and the study of cost efficiency
  - Estimation of revenue functions and revenue efficiency
  - This technique is also used in the estimation of multi-output and multi-input distance functions
  - Potential for applications in other disciplines
- Farrell (1957) poses that the production function is never well known suggesting the CD function while Aigner and Chu (1968) considered an estimation of it.

# Going back....

- Much of the work on stochastic frontiers began in 70's. Major contributions from Aigner, Schmidt, Lovell, Battese and Coelli and Kumbhakar.
- Ordinary least squares (OLS) regression production functions:
  - fit a function through the centre of the data
  - assumes all firms are efficient
- Deterministic production frontiers:
  - fit a frontier function over the data
  - assumes there is no data noise
- SFA production frontiers are a “mix” of these two

# Methodology

- Following Aigner and Chu (1968)

$$\ln(y_i) = x_i\beta - u_i$$

$x_i \in (K + 1)$  row vector,

$\beta \in (K + 1)$  column vector,

$u_i$  non negative error term

$$\text{Min } \sum_i u_i = \sum_i (\ln y_i - x_i\beta)$$

subject to  $u_i \geq 0$ .



- Thus,  $TE_i = \frac{y_i}{e^{x_i\beta}} = \frac{e^{x_i\beta - u_i}}{e^{x_i\beta}} = e^{-u_i}$  estimation using LP
- The specific measure is an output oriented Farrell measure of technical efficiency taking values between zero and one.
- It is important, however, to know the distribution of errors. Afriat (1972) assuming gamma estimates using a ML method, Richmond added Corrected Ordinary Least Squares while Schmidt (1976) assumes exponential or half normal random variables (If we apply OLS, intercept estimate is biased downwards, all other parameters are unbiased)

# Deterministic Frontier models

- So COLS suggests that the OLS estimator from OLS be corrected.
- If we do not wish to make use of any probability distribution for  $y_i$  then

$$\hat{\beta}_o(COLS) = \hat{\beta}_o(OLS) + \text{maximum}_i [\hat{u}_i : i = 1, 2, \dots, N]$$

where  $\hat{u}_i$  is the OLS residual for i-th firm.

- If we assume that  $u_i$  is distributed as Gamma then

$$\hat{\beta}_o(COLS) = \hat{\beta}_o(OLS) + \hat{\sigma}^2[OLS]$$

- It is a bit more complicated if  $u_i$  follows half-normal distribution.

# The Stochastic Production Function I

- It is a relationship between output and a set of input quantities.
  - We use this when we have a single output
  - In case of multiple outputs:
    - people often use revenue (adjusted for price differences) as an output measure
    - It is possible to use multi-output distance functions to study production technology.
- Aigner et al (1977) and Meeusen and van den Broeck (1977) proposed independently the functional relationship is usually written in the form:

$$\ln(y_i) = x_i\beta + v_i - u_i$$

$x_i \in (K + 1)$  row vector,

$\beta \in (K + 1)$  column vector,

$u_i$  non negative error term. *Therefore*  $\ln(y_i) \leq x_i\beta$

$v_i$  measurement errors

# The Stochastic Production Function II

- Crucial assumption by Aigner et al (1977)

$$\ln(y_i) = x_i\beta + v_i - u_i$$

$x_i \in (K + 1)$  row vector,

$\beta \in (K + 1)$  column vector,

$u_i$  non negative error term

$v_i$  measurement errors

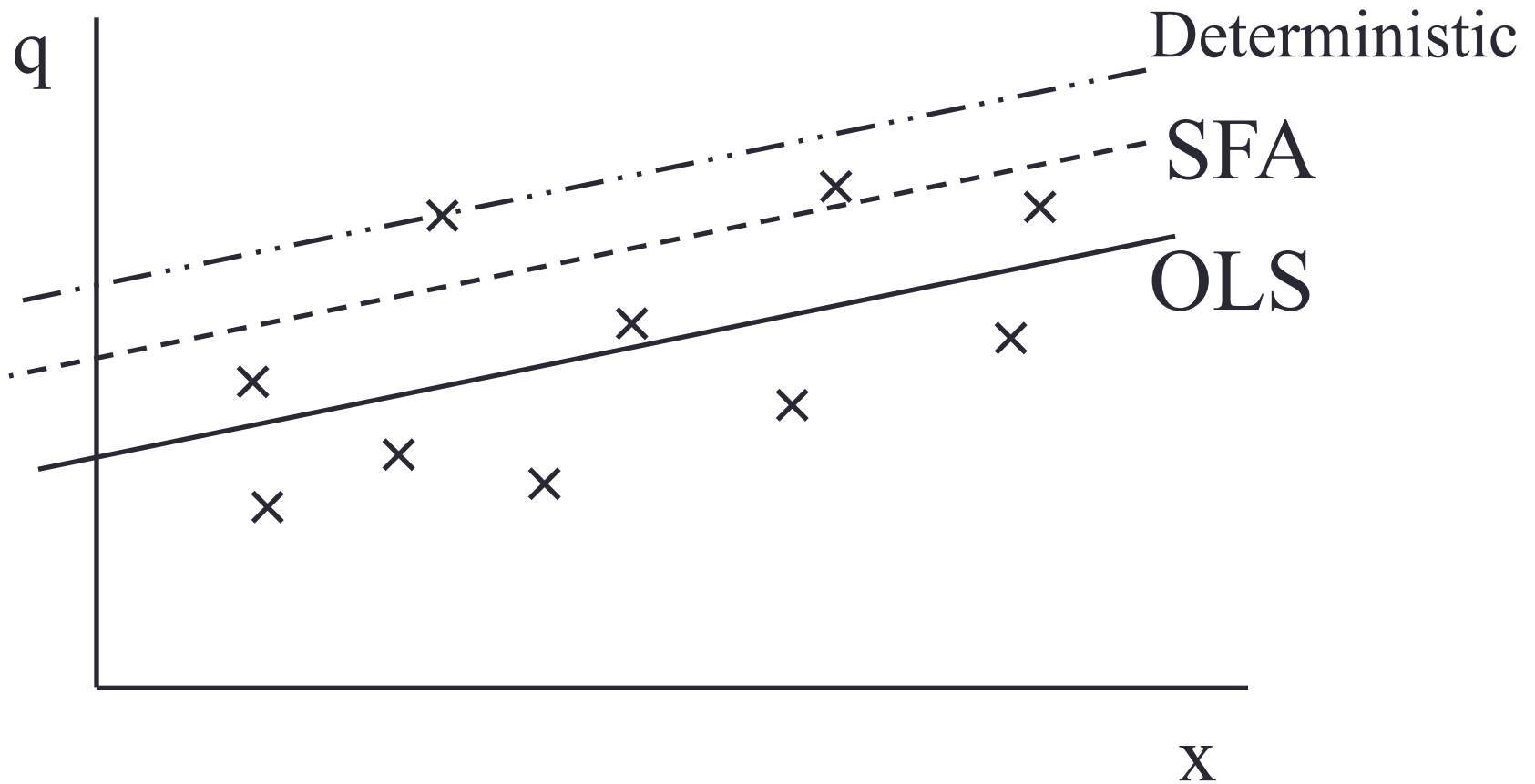
that

1.  $v_i$  = “noise” error term – symmetric i.i.d (eg. normal distribution  $v_i \sim (0, \sigma_v^2)$ )
2.  $u_i$  = “inefficiency error term” - non-negative (eg. half-normal distribution)

Do you think that these assumptions are correct? Criticism!!

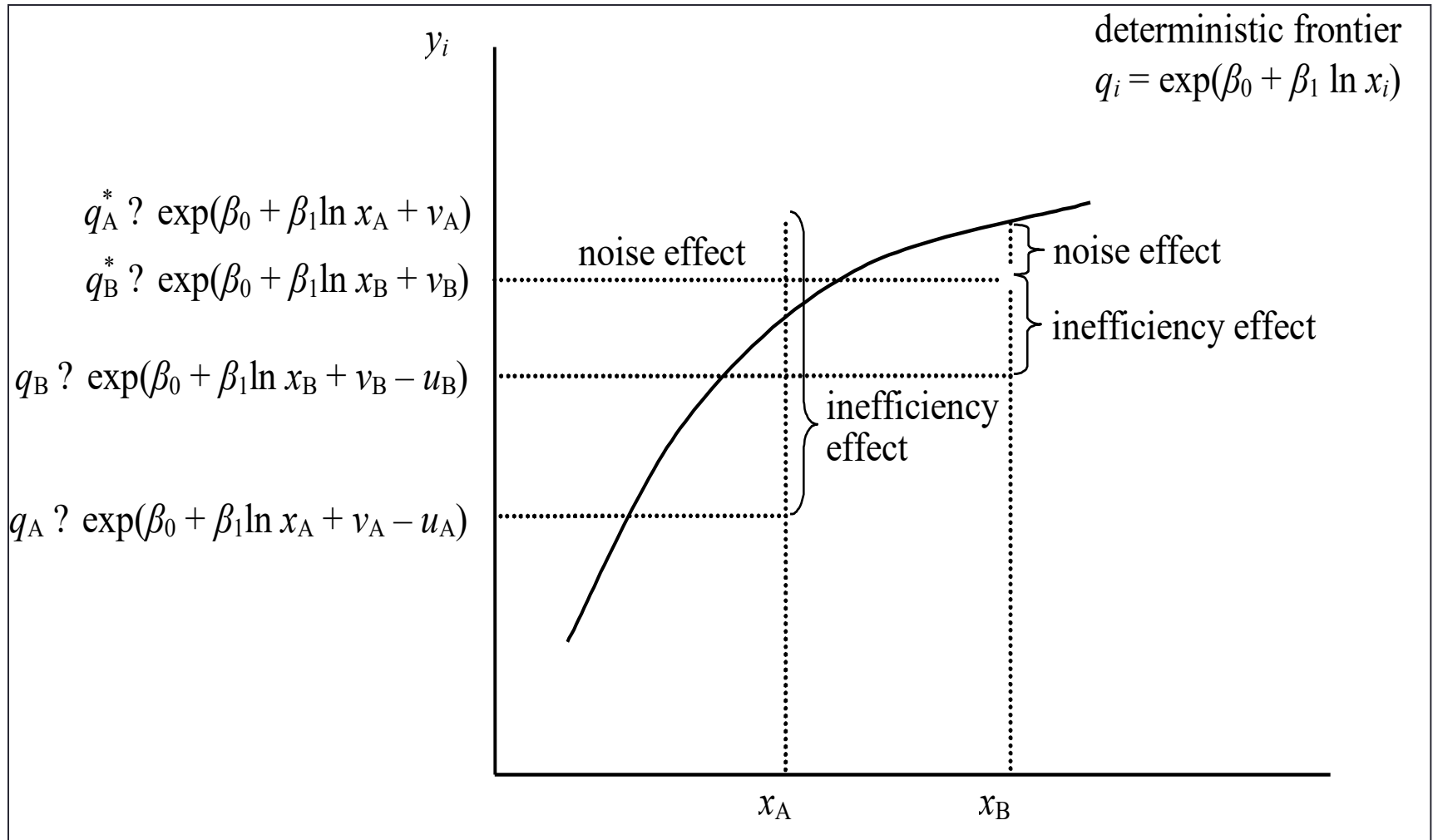
The model is called stochastic because the output values are bounded by the stochastic variable.

# Production functions/frontiers





# Stochastic frontiers-graphical representation



# Maximum Likelihood Estimation I

Let  $X_1, X_2, \dots, X_n$  be a random sample from a distribution that depends on one or more unknown parameters  $\theta_1, \theta_2, \dots, \theta_m$  with probability density (or mass) function  $f(x_i; \theta_1, \theta_2, \dots, \theta_m)$ . Suppose that  $(\theta_1, \theta_2, \dots, \theta_m)$  is restricted to a given parameter space  $\Omega$ . When regarded as a function of  $\theta_1, \theta_2, \dots, \theta_m$ , the joint probability density (or mass) function of  $X_1, X_2, \dots, X_n$ :

$$L(\theta_1, \theta_2, \dots, \theta_m) = \prod_{i=1}^n f(x_i; \theta_1, \theta_2, \dots, \theta_m)$$

$((\theta_1, \theta_2, \dots, \theta_m) \text{ in } \Omega)$  is called the **likelihood function**. If now

$$\left[ u_1(x_1, x_2, \dots, x_n), u_2(x_1, x_2, \dots, x_n), \dots, u_m(x_1, x_2, \dots, x_n) \right]$$

is the  $m$ -tuple that maximizes the likelihood function, then

$\theta_i \left[ u_i(X_1, X_2, \dots, X_n) \right]$  is the **maximum likelihood estimator** of  $\theta_i$ , for  $i = 1, 2, \dots, m$ . The corresponding observed values of the statistics in

$$\left[ u_1(x_1, x_2, \dots, x_n), u_2(x_1, x_2, \dots, x_n), \dots, u_m(x_1, x_2, \dots, x_n) \right]$$

are called the maximum likelihood estimates of  $\theta_i$ , for  $i = 1, 2, \dots, m$ .

# Maximum Likelihood Estimation II

Let us now have a look at some well-known distributions:

$$f(x_i; \theta) = \frac{1}{2\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\theta)^2}{2\sigma^2}} \Leftrightarrow L(\theta) = \prod_{i=1}^n f(x_i; \theta) = \frac{1}{2\sigma^n (2\pi)^{n/2}} e^{-\frac{\sum_{i=1}^n (x_i-\theta)^2}{2\sigma^2}} \Leftrightarrow$$

$$\log L(\theta) = -\frac{n}{2} \log \sigma^2 - \frac{n}{2} \log(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta)^2 \Leftrightarrow \frac{\partial \log L(\theta)}{\partial \theta} =$$

$$f(x_i; \theta) = e^{-\theta} \frac{\theta^x}{x!} \Leftrightarrow L(\theta) = \prod_{i=1}^n f(x_i; \theta) = e^{-n\theta} \frac{\theta^{\sum_{i=1}^n x_i}}{x_1! \dots x_n!} \Leftrightarrow$$

$$\log L(\theta) = -n\theta + \theta \sum_{i=1}^n x_i - \log(x_1! \dots x_n!) \Leftrightarrow \frac{\partial \log L(\theta)}{\partial \theta} =$$

$$f(x_i; \theta) = \frac{1}{\theta} \Leftrightarrow L(\theta) = \prod_{i=1}^n f(x_i; \theta) = \begin{cases} \frac{1}{\theta^n}, 0 \leq x_i \leq \theta \\ 0, \alpha\lambda\lambda\iota\omega\varsigma \end{cases} = \begin{cases} \frac{1}{\theta^n}, \theta = \max(x_1, x_2, \dots, x_n) \\ 0, \alpha\lambda\lambda\iota\omega\varsigma \end{cases}$$

$$f(x_i; \theta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$$

# Maximum Likelihood Estimation III

- The parameters of the stochastic frontier production function can be estimated using the ML method (or COLS by Richmond (1974)).

- Parameters to be estimated in a standard SF model are:

$$\beta, \sigma_v^2 \text{ and } \sigma_u^2$$

- Likelihood methods are used in estimating the unknown parameters. Coelli (1995)'s Monte Carlo study shows that in large samples MLE is better than COLS.
- According to Aigner et al., (1977) expressed the likelihood function in terms of  $\sigma_s^2 \equiv \sigma_u^2 + \sigma_v^2, \lambda \equiv \frac{\sigma_u}{\sigma_v}$
- While Battese and Corra (1977) shows that while testing for the presence of technical inefficiency depends upon the parametrization used  $\gamma \equiv \frac{\sigma_u^2}{\sigma_s^2} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$

# Mean Technical Efficiency

- Battese and Corra (1977) shows that in terms of parameterization the log-likelihood function is equal to:

$$\ln(L) = \frac{-N}{2} \ln\left(\frac{\pi}{2}\right) - \frac{N}{2} \log(\sigma_s^2) + \sum_{i=1}^N \ln[1 - \Phi(z_i)] - \frac{1}{2\sigma_s^2} \sum_{i=1}^N (\ln y_i - x_i \beta)^2$$

$$z_i = \frac{(\ln y_i - x_i \beta)}{\sigma_s} \sqrt{\frac{\gamma}{1-\gamma}},$$

$\Phi(\bullet)$  distribution function of the standard normal random variable

Mean Technical Efficiency

$\beta$ ,  $\sigma_v^2$  and  $\sigma_u^2$

$$\rightarrow E(e^{-u_i}) = 2 \left[ \left[ 1 - \Phi\left(\sigma_s \sqrt{\gamma}\right) \right] e^{-\gamma \sigma_s^2 / 2} \right]$$

# Production Function Specification

- A number of different functional forms are used in the literature to model production functions:
  - Cobb-Douglas (linear logs of outputs and inputs)
  - Quadratic (in inputs)
  - Normalised quadratic
  - Translog function

$$\ln q = \beta_0 + \sum_{n=1}^N \beta_n \ln x_n + \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \beta_{nm} \ln x_n \ln x_m + u$$

- Translog function is very commonly used – it is a generalisation of the Cobb-Douglas function
- It is a flexible functional form providing a second order approximation
- Cobb-Douglas and Translog functions are linear in parameters and can be estimated using least squares methods.
- It is possible to impose restrictions on the parameters (homogeneity conditions)

# Cobb-Douglas Functional form

- $\ln y_i = \ln q_i = \beta_0 + \beta_1 \ln x_{1i} + \beta_2 \ln x_{2i} + v_i - u_i$
- Linear in logs
- Advantages:
  - easy to estimate and interpret
  - requires estimation of few parameters:  $K+3$
- Disadvantages:
  - simplistic - assumes all firms have same production elasticities and that substitution elasticities equal 1

# Translog Functional form

- $\ln q_i = \beta_0 + \beta_1 \ln x_{1i} + \beta_2 \ln x_{2i} + 0.5\beta_{11}(\ln x_{1i})^2 + 0.5\beta_{22}(\ln x_{2i})^2 + \beta_{12} \ln x_{1i} \ln x_{2i} + v_i - u_i$
- Quadratic in logs
- Advantages:
  - flexible functional form - less restrictions on production elasticities and substitution elasticity
- Disadvantages:
  - more difficult to interpret
  - requires estimation of many parameters:  $K+3+K(K+1)/2$
  - can suffer from curvature violations



# Interpretation of estimated parameters

Cobb-Douglas:

Production elasticity for j-th input is:  $E_j = \beta_j$

Scale elasticity is:  $\varepsilon = E_1 + E_2$

Translog:

Production elasticity for i-th firm and j-th input is:  $E_{ji} = \beta_j + \beta_{j1} \ln x_{1i} + \beta_{j2} \ln x_{2i}$

Scale elasticity for i-th firm is:  $\varepsilon_i = E_{1i} + E_{2i}$

Note: If we use transformed data where inputs are measured relative to their means, then Translog elasticities at means would simply be  $\beta_i$ .

# Tests of hypotheses I

e.g., Is there significant technical inefficiency?

$H_0: \gamma=0$  versus  $H_1: \gamma>0$

Value  $\gamma=0$  denotes that the deviation from the frontier is due entirely to noise while  $\gamma=1$ , represent that all deviation are due to technical efficiency. The previous specification allows us to examine the null hypothesis that they are not technical efficiency effects in the model versus the alternative hypothesis .

Test options:

- t-test

$$\text{t-ratio} = (\text{parameter estimate}) / (\text{standard error})$$

- Likelihood ratio (LR) test

[note that the above hypothesis is one-sided - therefore must use Kodde and Palm critical values (not chi-square) for LR test

- LR test “safer”

# Likelihood ratio (LR) tests

## Steps:

- 1) Estimate unrestricted model ( $LLF_1$ )
- 2) Estimate restricted model ( $LLF_0$ )  
(eg. set  $\gamma=0$ )
- 3) Calculate  $LR=-2(LLF_0-LLF_1)$
- 4) Reject  $H_0$  if  $LR > \chi_R^2$  table value,  
where  $R$  = number of restrictions  
(Note: Kodde and Palm tables must be used if test is one-sided)

# Tests of hypotheses II

As concerns the nature of technical efficiency, the stochastic frontier model is well- defined by the following three subcases,

- i) when  $\gamma = \delta_i = 0$  , there is no technical inefficiency deterministic or stochastic,
- ii) when  $\gamma = 0$  , where there is deterministic
- iii) when all  $\delta_i = 0$  parameters (except  $\delta_0$  ) are zero and the variables do not affect technical efficiency levels and the model reduces to the one proposed by Stevenson (1980).

The nature of technical inefficiencies can be examined by conducted a null hypothesis of  $\lambda = 0$  versus  $\lambda > 0$  the alternative of can be tested by using the well known generalized likelihood ratio statistic.

# Test for Cobb-Douglas versus Translog

- Using sample data file which comes with the FRONTIER program
- $H_0: \beta_{11}=\beta_{22}=\beta_{12}=0$ ,  $H_1: H_0$  false
- Compute  $-2[LLF_0-LLF_1]$  which is distributed as Chi-square ( $r$ ) under  $H_0$ .
- For example, if:  
 $LLF_1=-14.43$ ,  $LLF_0=-17.03$   
 $LR=-2[-17.03-(-14.43)]=5.20$   
Since  $\chi_3^2$  5% table value = 7.81  $\Rightarrow$  do not reject  $H_0$

# Stochastic Frontier: Model Specification

$$q_i = \underbrace{\exp(\beta_0 + \beta_1 \ln x_i)}_{\text{deterministic component}} \times \underbrace{\exp(v_i)}_{\text{noise}} \times \underbrace{\exp(-u_i)}_{\text{inefficiency}}$$

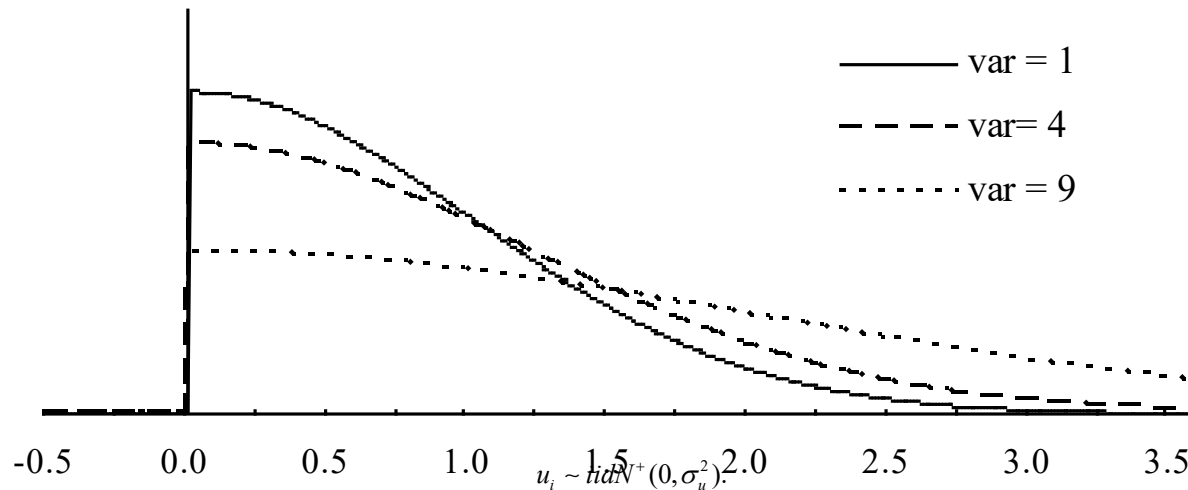
In general, we write the stochastic frontier model with several inputs and a general functional form (which is linear in parameters) as

$$\ln q_i = \mathbf{x}'_i \boldsymbol{\beta} + v_i - u_i$$

- We stipulate that  $u_i$  is a non-negative random variable
- By construction the inefficiency term is always between 0 and 1.
- This means that if a firm is inefficient, then it produces less than what is expected from the inputs used by the firm at the given technology.
- We can define technical efficiency as the ratio of “observed” or “realised output” to the stochastic frontier output

$$TE_i = \frac{q_i}{\exp(\mathbf{x}'_i \boldsymbol{\beta} + v_i)} = \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + v_i - u_i)}{\exp(\mathbf{x}'_i \boldsymbol{\beta} + v_i)} = \exp(-u_i)$$

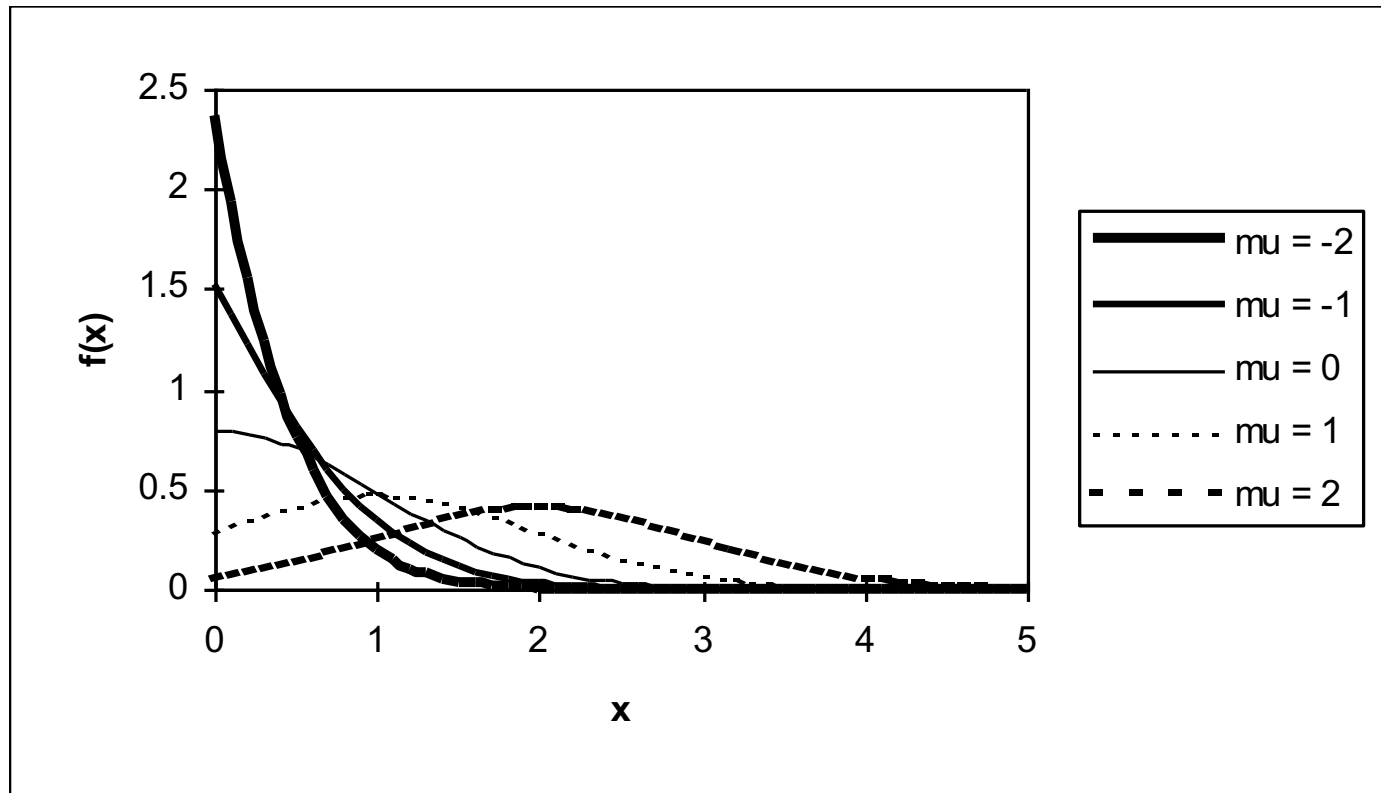
# Truncated normal distribution for u



## Distribution of u:

We note that: As u is truncated from a normal distribution with mean equal to 0,  $E(u)$  is “towards” zero and therefore technical efficiency tends to be high just by model construction.

# Truncated normal with non-zero means



A more general specification:  $u \sim N^+(\mu, \sigma_u^2)$

This forms the basis for the inefficiency effects model where

$$u \sim N^+(\mu_i, \sigma_u^2)$$

$$\mu_i = \delta_0 + \sum_k \delta_k Z_{ki}$$



# Energy Efficiency Greek Firms

## Data:

- 161 Greek firm that invest in Energy Efficiency 60% response rate – response bias
- Data validation
- Actual observations used: 1499
- The deflated total value of shipments in thousands euros (based year 1990)
- input variables the labor of each firms based on annual full-time equivalents and the deflated total value of assets in thousands euros.

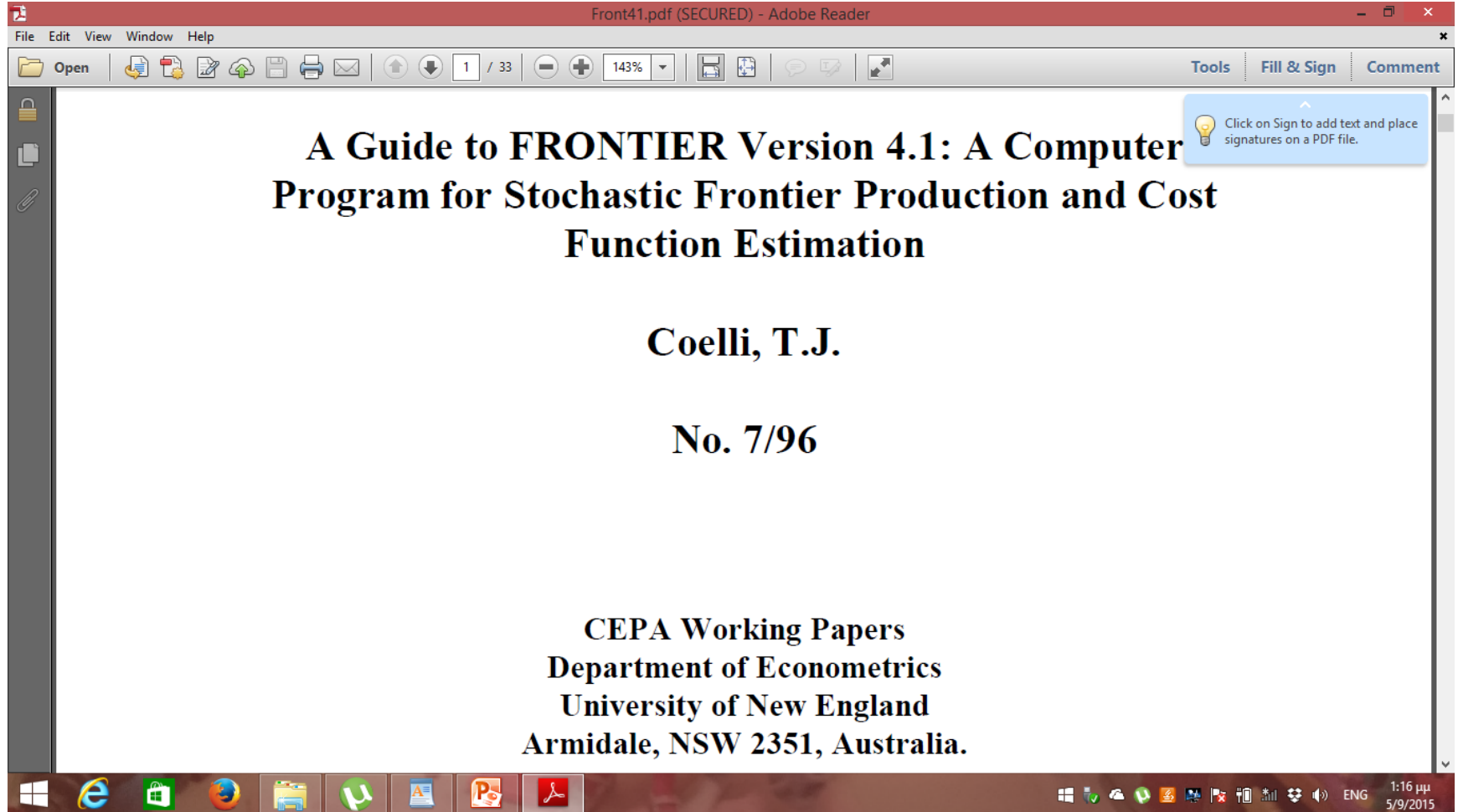
## Methods:

DEA and SFA Methods

# Estimation of SFA Models

- In the case of translog model, it is a good idea to transform the data – divide each observation by its mean
  - Then the coefficients of  $\ln X_i$  can be interpreted as elasticities.
- Most standard packages such as STATA and LIMDEP.
- FRONTIER by Coelli is a specialised program for purposes of estimating SF models.
  - Available for free downloads from CEPA website:  
[www.uq.edu.au/economics/cepa](http://www.uq.edu.au/economics/cepa)

# FRONTIER Program I



# FRONTIER Program II

- We need:
  1. FRONTIER.EXE file
  2. FRONT41.000 start up file
  3. Data file (\*.DAT)
  4. INSTRUCTION FILE (\*.INS)
  5. Finally we get the output file (\*.OUT)

# FRONTIER Program III

Data file (\*.DAT) structure.

1. Firm number
2. Period Number
3. Output Y
4. Input X
5. ..
6. Input X's
7. Environmental variables.

# FRONTIER Instruction File

**Table** The FRONTIER Instruction File

1	1=ERROR COMPONENTS MODEL, 2=TE EFFECTS MODEL
chap9.txt	DATA FILE NAME
chap9_2.out	OUTPUT FILE NAME
1	1=PRODUCTION FUNCTION, 2=COST FUNCTION
y	LOGGED DEPENDENT VARIABLE (Y/N)
344	NUMBER OF CROSS-SECTIONS
1	NUMBER OF TIME PERIODS
344	NUMBER OF OBSERVATIONS IN TOTAL
10	NUMBER OF REGRESSOR VARIABLES (Xs)
n	MU (Y/N) [OR DELTA0 (Y/N) IF USING TE EFFECTS MODEL]
n	ETA (Y/N) [OR NUMBER OF TE EFFECTS REGRESSORS (Zs)]
n	STARTING VALUES (Y/N)

- Here MU refers to inefficiency effects models and ETA refers to time-varying inefficiency effects (we will come to this shortly)
- The program uses the ratio of variances as the transformation
- It allows for the use of single cross-sections as well as panel data sets

# FRONTIER output

the final mle estimates are :

	coefficient	standard-error	t-ratio
beta 0	-0.90316475E+00	0.16498610E+00	-0.54741868E+01
beta 1	0.71424809E+00	0.33469693E+00	0.21340144E+01
beta 2	0.61516795E+00	0.69233590E+00	0.88853972E+00
beta 3	-0.47487501E+01	0.65654976E+01	-0.72328869E+00
beta 4	-0.92060716E-01	0.16858623E+00	-0.54607494E+00
beta 5	0.41198757E-01	0.34831979E+00	0.11827854E+00
beta 6	0.36624071E+02	0.40449416E+02	0.90542890E+00
beta 7	0.17381902E-01	0.82204147E-02	0.21144800E+01
beta 8	-0.16425365E+00	0.47591966E+00	-0.34512895E+00
beta 9	-0.17836537E+00	0.61338415E+00	-0.29078901E+00
beta10	0.12017607E-01	0.21218532E-01	0.56637317E+00
beta11	0.20892935E-03	0.11544911E-02	0.18097095E+00
beta12	-0.14792861E-02	0.38047266E-02	-0.38880221E+00
beta13	-0.20050694E-02	0.51137679E-02	-0.39209238E+00
beta14	0.81014619E-01	0.13583193E+00	0.59643280E+00
beta15	0.90697900E-01	0.69979449E-01	0.12960648E+01
sigma-squared	0.10624287E+01	0.10111001E+00	0.10507651E+02
gamma	0.41538781E+00	0.57926074E-01	0.71709988E+01
mu is restricted to be zero			
eta is restricted to be zero			

log likelihood function = -0.18425896E+04

LR test of the one-sided error = 0.12558383E+03  
with number of restrictions = 1

# SF Models - continued

Predicting Firm Level Efficiencies:

Once the SF model is estimated using MLE method, we compute the following:

$$u_i^* = -(\ln q_i - \mathbf{x}_i' \boldsymbol{\beta}) \sigma_u^2 / \sigma^2 \quad \text{and} \quad \sigma_*^2 = \sigma_v^2 \sigma_u^2 / \sigma^2.$$

We use estimates of unknown parameters in these equations and compute the best predictor of technical efficiency for each firm  $i$  :

$$TE_i \hat{=} E \left\{ \exp(-u_i) \mid q_i \right\} = \left[ \Phi \left( \frac{u_i^*}{\sigma_*} - \sigma_* \right) \right] / \left[ \Phi \left( \frac{u_i^*}{\sigma_*} \right) \right] \exp \left\{ \frac{\sigma_*^2}{2} - u_i^* \right\}.$$

We use standard normal density and distribution functions to evaluate technical efficiency.



# SF Models - continued

Industry efficiency:

- Industry efficiency can be computed as the average of technical efficiencies of the firms in the sample
- Industry efficiency can be seen as the expected value of a randomly selected firm from the industry. Then we have

$$TE \hat{E} \equiv E \{ \exp(-u_i) \} = 2\Phi(-\sigma_u) \exp \left\{ \frac{\sigma_u^2}{2} \right\}.$$

Confidence intervals for technical efficiency scores (for the firms and the industry as a whole) can also be computed.

- We note that there are no firms with a TE score of 1 as in the case of DEA.
- No concept of peers exists in the case of SFA.

# FRONTIER output continued

technical efficiency estimates :

firm	eff.-est.	
1	0.77532384	
2	0.72892751	
3	0.77332991	
341	0.76900626	-
342	0.92610064	
343	0.81931012	
344	0.89042718	
mean efficiency =		0.72941885

- **Mean efficiency can be interpreted as the “industry efficiency”.**

- Example - estimate translog production function using sample data file which comes with the FRONTIER program - 151 firms
- t-ratio for  $\gamma = 24.36$ , and  $N(0,1)$  critical value at 5% = 1.645  $\Rightarrow$  reject  $H_0$
- Or the LR statistic = 28.874, and Kodde and Palm critical value at 5% = 2.71  $\Rightarrow$  reject  $H_0$

The LR statistic has mixed Chi-square distribution

# Distributional assumptions – the truncated normal distribution

- $N(\mu, \sigma^2)$  truncated at zero
- More general patterns
- Can test hypothesis that  $\mu=0$  using t-test or LR test
- The restriction  $\mu=0$  produces the half-normal distribution:  
 $|N(0, \sigma^2)|$

# Scale efficiency

- For a Translog Production Function (Ray, 1998)
- An output-orientated scale efficiency measure is:

$$SE_i = \exp[(1-\varepsilon_i)^2/2\beta]$$

where  $\varepsilon_i$  is the scale elasticity of the  $i$ -th firm and

$$\beta = \sum_{j=1}^K \sum_{k=1}^K \beta_{jk}$$

- If the frontier is concave in inputs then  $\beta < 0$ . Then SE is in the range 0 to 1.

# Stochastic Frontier Models: Some Comments

We note the following points with respect to SFA models

- It is important to check the regularity conditions associated with the estimated functions – local and global properties
  - This may require the use of Bayesian approach to impose inequality restrictions required to impose convexity and concavity conditions.
- We need to estimate distance functions directly in the case of multi-output and multi-input production functions.
- It is possible to estimate scale efficiency in the case of translog and Cobb-Douglas specifications

# Panel data models

- Data on  $N$  firms over  $T$  time periods
- Investigate technical efficiency change (TEC)
- Investigate technical change (TC)
- More data = better quality estimates
- Less chance of a one-off event (eg. climatic) influencing results
- Can use standard panel data models
  - no need to make distributional assumption
  - but must assume TE fixed over time
- The model:  $i=1,2,\dots,N$  (cross-section of firms);  $t=1,2,\dots,T$  (time points)

$$\ln y_{it} = x_{it}\beta + v_{it} - u_{it}; v_{it} \approx N(0, \sigma_v^2); u_{it} \approx N^+(0, \sigma_u^2)$$

# Panel data models

Some Special cases:

1. Firm specific effects are time invariant:  $u_{it} = u_i$ .
2. Time varying effects: Kumbhakar (1990)

$$u_{it} = \left[1 + \exp(bt + ct^2)\right]^{-1} u_i$$

3. Time-varying effects with convergence – Battese and Coelli (1992)

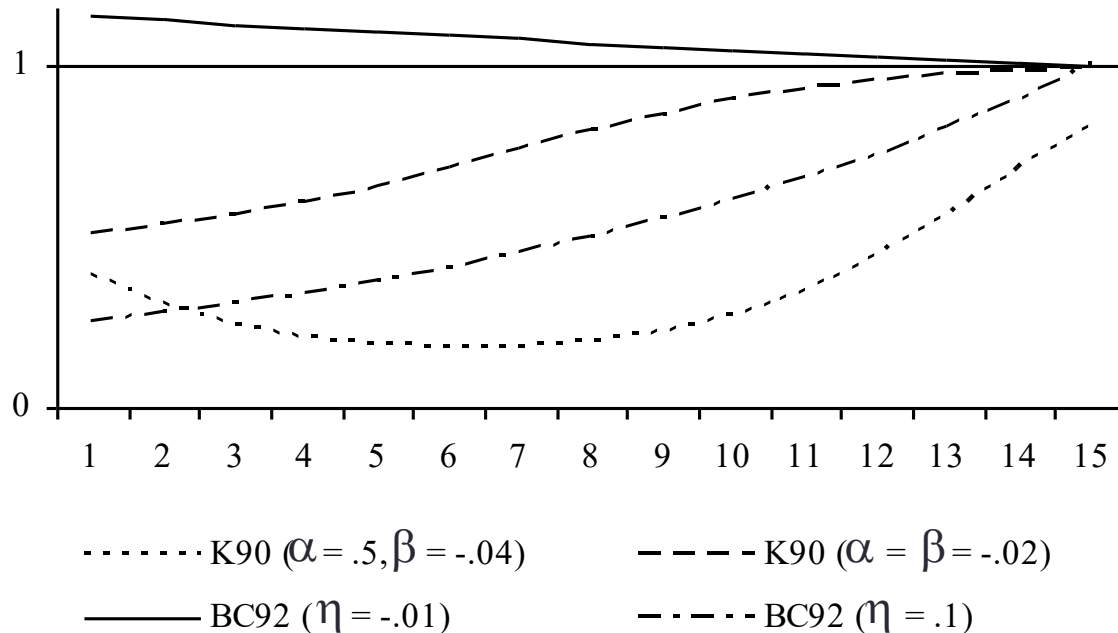
$$u_{it} = \left[\exp\{-\eta(t - T)\}\right] u_i$$

Sign of  $\eta$  is important. As  $t$  goes to  $T$ ,  $u_{it}$  goes to  $u_i$ .

In FRONTIER Program, this is under Error Components Model.



# Time profiles of efficiencies



**Note:** These are all smooth functions of trends of technical efficiency over time. These trends are also independent of any other data on the firms. There is scope for further work in this area.

# Accounting for Production Environment (Battese & Coelli model 1995)

- Technical efficiency is influenced by exogenous factors that characterise the environment in which production takes place
  - Government regulation, ownership, education level of the farmer, etc.
- Non-stochastic Environmental Variables

In this case firm-level technical efficiency levels predicted will vary with traditional inputs and environmental variables.

$$\ln q_i = x_i \beta + z_i \gamma + v_i - u_i$$

- Inefficiency effects model (Battese, Coelli 1995)

$$\ln Y_i = \alpha_0 + \sum_{i=K,L} \alpha_i \ln X_i + \frac{1}{2} \sum_{i=K,L} \sum_{j=K,L} \alpha_{ij} \ln X_i \ln X_j + \alpha_t t + \frac{1}{2} \alpha_{tt} t^2 + \sum_{i=K,L} \alpha_{it} t \ln X_i + u_i - v_i$$

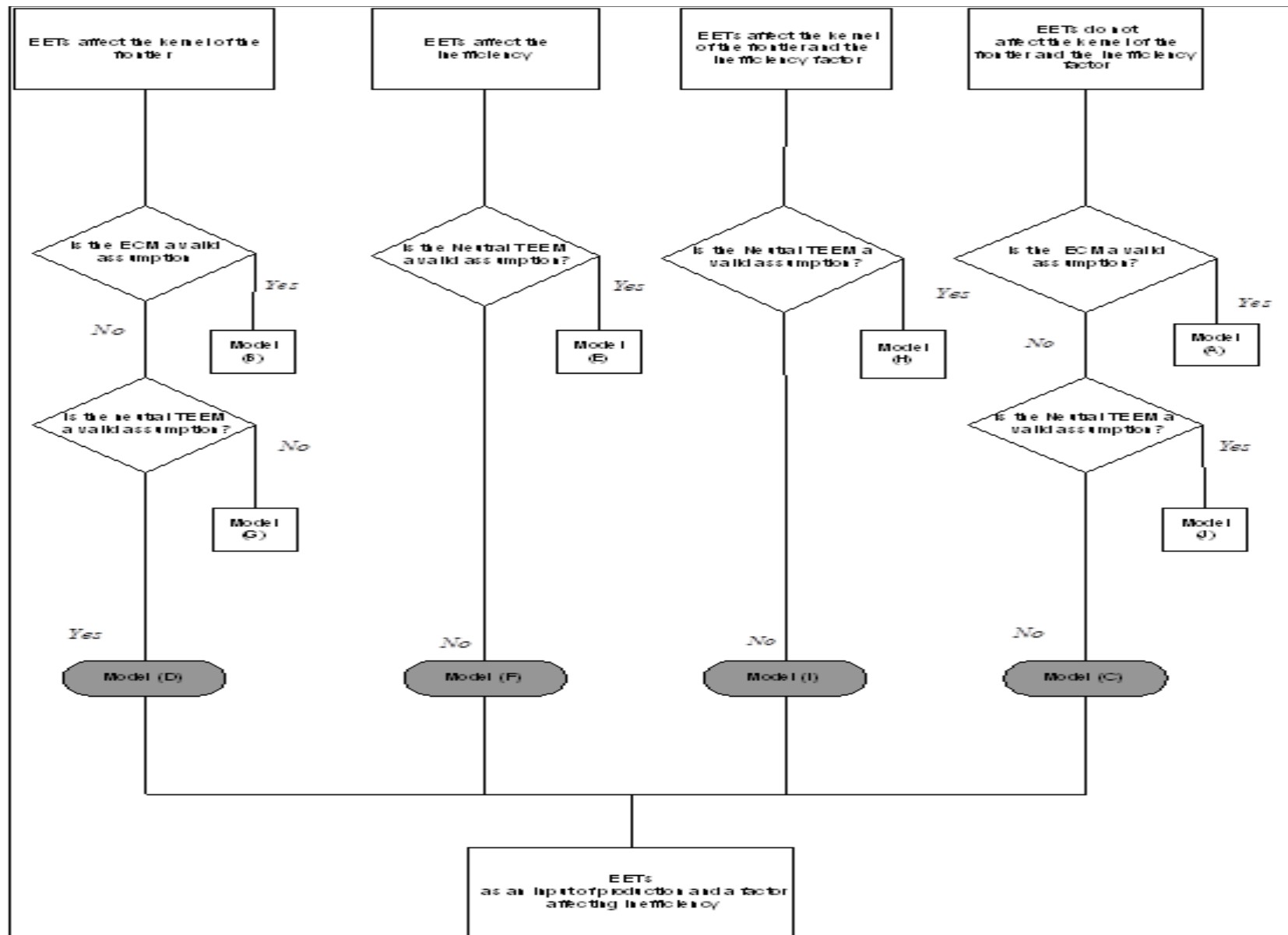
where  $\delta$  is a vector of parameters to be estimated. In the FRONTIER program, this is the TEEFFECTS model

$$v_i = \delta_0 + \sum_{j=1}^8 \delta_j z_j + \sum_{i=1}^8 \sum_{k=1}^3 \gamma_{kj} z_{ji} x_{ki} + w_i$$

## Current research

- We have seen how technical efficiency can be computed, but it is difficult to compute standard errors.
- Peter Schmidt and his colleagues have been working on a number of related topics here.
  - Bootstrap estimators and confidence intervals for efficiency levels in SF models with panel data
  - Testing whether technical inefficiency depends on firm characteristics
  - On the distribution of inefficiency effects under different assumptions
- Bayesian estimation of stochastic frontier models
  - Posterior distribution of technical efficiencies
  - Estimation of distance functions

Source: Kounetas, K., and Tsekouras, K., 2010. Are the Energy Efficiency Technologies Efficient? Economic Modeling 29, 1798-1808



# FRONTIER IN R

- `install.packages ( "frontier" )`
- `data( front41Data )`
- `sfaResult <- sfa( log( output ) ~ log( capital ) + log( labour ),`
- `data = (front41Data )`
- `coef( summary( sfaResult ), which = "ols" )`
- `coef( summary( sfaResult ) )`
- `coef( summary( sfaResult, extraPar = TRUE ) )`

# REFERENCES I

- Battese, G. E., 1997. A Note on the Estimation of Cobb-Douglas Production Functions when Some Explanatory Variables have Zero Values. *Journal of Agricultural Economics* 48, 250-252.
- Battese, G.E., Broca, S.S., 1997. Functional forms of stochastic frontier production functions and models for technical inefficiency effects: a comparative study for wheat farmers in Pakistan. *Journal of Productivity Analysis* 8, 395-414.
- Battese, G. E., Coelli, T., 1992. Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis* 3 (12), 153-69.
- Battese, G.E. Coelli, T., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20, 325-32.
- Coelli, Tim J., 1996. A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation. Armidale, NSW, Australia: Department of Econometrics, University of New England.

# REFERENCES II

- Coelli, T.J. Prasada D. R. and G. E. Battese. 2005. An Introduction to Efficiency and Productivity Analysis. Springer: New York.
- Farrell, M.J., 1957. The Measurement of Productive Efficiency. Journal of the Royal Statistical Society Series A, 1957, 120, 253-90.
- Huang, C.J. Liu, J.T., 1994. Estimation of a non neutral stochastic frontier function, Journal of Productivity Analysis 15, 171-180.
- Kounetas, K., and Tsekouras, K., 2010. Are the Energy Efficiency Technologies Efficient? Economic Modeling 29, 1798-1808
- Kodde, D.A., Palm, F.C., 1986. Wald criteria for jointly testing equality and inequality restrictions, Econometrica 54, 1243-1248.
- Kumbhakar S. C., Knox Lovell, C. A., 2000. Stochastic Frontier Analysis, Cambridge University Press.
- <https://www.youtube.com/watch?v=fvNUUJuFXM0>