**Cost, Revenue, and Profit Function Estimates**

**Levent Kutlu**

**University of Texas at Arlington**


**Shasha Liu**

**Rice University**


**Robin C. Sickles**

**Rice University**

This revision

October 15, 2018

# 1. Introduction

The purpose of this chapter is to provide a review of how cost, revenue, and profit functions are used to identify and characterize an underlying technology. Such an undertaking for a Handbook will undoubtedly leave out certain topics. We will provide a relative cursory discussion of duality theory and the links between cost, revenue, and profit functions and the underlying technology they characterize under certain testable regularity conditions. A more extensive recent treatment and summary can be found in Sickles and Zelenyuk (2018).

Moreover, as the functional forms and estimation set-up for the cost, revenue, and profit functions have many generic commonalities, we will concentrate on the more widely used functions to motivate various issues in the flexibility of various parametric functions, in the imposition of regularity conditions, in the use of non-parametric estimation of models, and in standard econometric models used to estimate the parameters of these different functional characterizations of an underlying technology.

We also discuss briefly modeling settings in which allocative and technical distortions may exist and how such distortions may be addressed empirically in the specification and estimation of the dual functional representations of the underlying primal technology.

# 2. Duality of the Technology and Characterizations of the Technology using the Cost, Revenue, and Profit Functions[1]

Very often researchers either do not have information that allows them to identify the underlying technology and thus its characterization in terms of marginal products, substitution possibilities, and other technical aspects of the production process, or have problems estimating such a relationship due to statistical problems such as endogeneity of inputs. This situation was one of the motivations for work in the area of duality by various legendary economists. Among the masterminds, Ronald Shephard revolutionized the neoclassical production theory by developing his duality theory, a foundation for many practical results later on. We will summarize and highlight some important results of this theory that we will utilize in our

---

[1] For more details on the issues discussed in this section see chapter 2 of Sickles and Zelenyuk (2018) whose notation we adopt here.

discussion of various estimating relationships that rest on this theory, such as the cost, profit, and revenue functions that are the topics of this Chapter.

The starting point is a firm that produces $M_y$ outputs $\in \mathfrak{R}_+^{M_y}$, using $M_x$ inputs $x \in \mathfrak{R}_+^{M_x}$ with exogenous prices $= (w_1, \ldots, w_{M_x}) \in \mathfrak{R}_{++}^{M_x}$, using some technology $T$, where the technology set $T$ is defined as $T \equiv \{(x, y) \in \mathfrak{R}_+^{N_x} \times \mathfrak{R}_+^{M_y} : y \text{ is producible from } x\}$. The input requirement set $L(y)$ completely characterizes the technology and is defined as

$$L(y) \equiv \{x \in \mathfrak{R}_+^{M_x} : y \text{ is producible from } x\}, \quad y \in \mathfrak{R}_+^{M_y}. \tag{2.1}$$

Moreover, the Shephard's input distance function, defined as

$$D_i(y, x) \equiv \sup\{\theta > 0 : x/\theta \in L(y)\} \tag{2.2}$$

completely characterizes the input requirement set in the sense

$$x \in L(y) \quad \Leftrightarrow \quad D_i(y, x) \geq 1. \tag{2.3}$$

A firm faced with a cost constraint chooses its level of inputs given the price and output level. Such a cost (we are considering here long-run costs) and its functional representation can be shown to be

$$C(y, w) \equiv \min_x \{wx : x \in L(y)\}. \tag{2.4}$$

Given cost minimizing decision by the firm in the employment of its resources, we can express the input demand functions as

$$x(y, w) \equiv \operatorname*{argmin}_x \{wx : x \in L(y)\}, \tag{2.5}$$

which, of course, are conditional on the level of output produced. If the input requirement sets are convex and there is free disposability of inputs, then it can be shown that the technology underlying the cost function can be identified. Thus, we say that under these conditions the cost function is dual to the primal technology.

If a firm's behavioral objective is to maximize revenues instead of minimizing costs, then a duality can be shown to exist between the revenue function and the underlying primal technology under certain regularity conditions. First, define the technological possibilities

(output set) as $P(x) \equiv \{y \in \mathfrak{R}_+^{M_y}: y$ is producible from $x\}$, $x \in \mathfrak{R}_+^{M_x}$, and let the output prices

for the M outputs be $(p_1, \ldots, p_{M_y}) \in \mathfrak{R}_{++}^{M_y}$. The output set $P(x)$ completely characterizes

technology. The Shephard's output distance function is used to completely characterize $P(x)$ as

$$y \in P(x) \quad \Leftrightarrow \quad D_o(x,y) \leq 1, \tag{2.6}$$

where

$$D_o(x,y) \equiv \inf\{\theta > 0 : y/\theta \in P(x)\}. \tag{2.7}$$

The revenue function $R: \mathfrak{R}_+^{M_x} \times \mathfrak{R}_{++}^{M_y} \rightarrow \mathfrak{R}_+ \cup \{+\infty\}$, is then defined as

$$R(x,p) \equiv \max_y \{py : y \in P(x)\}, \tag{2.8}$$

which leads to a set of output supply functions

$$y(x,p) \equiv \underset{y}{\arg\max} \{py: y \in P(x)\}. \tag{2.9}$$

Finally, the profit function $\pi: \mathfrak{R}_{++}^{M_x} \times \mathfrak{R}_{++}^{M_y} \rightarrow \mathfrak{R}_+ \cup \{+\infty\}$ is defined as

$$\pi(w,p) \equiv \underset{x,y}{\sup}\{py - wx : (x,y) \in T\},$$

(2.10)

and the corresponding output supply and input demand equations are given by

$$(x(w,p), y(w,p)) \equiv \arg \underset{x,y}{\sup}\{py - wx: (x,y) \in T\}, \tag{2.11}$$

assuming profit maximizing behaviors.

Next, we turn to explicit functional forms and assumptions for cost functions and factor demand equations, revenue functions and output supply equations, and profit functions and the corresponding output supply and input demand equations. We also consider their shadow prices when allocative distortions exist in the optimal relative output mix and input mix selected by the firm.

## 3.    Cost Functions

Simple inflexible cost functions, thanks to their parametric forms, often satisfy the regularity conditions required in the production theory and dual forms such as the cost functions.

However, these simple but inflexible forms have serious limitations. Aside from the strong and often unrealistic restrictions they impose on the technology being modeled, they suffer from other shortcomings as well. A multi-output Cobb-Douglas distance function, for example, does not satisfy the concavity condition because it has a convex production possibility frontier. Since a majority of firms produce more than one output, distinguishing each output by using a different production function is empirically infeasible and theoretically dubious. Given the fact that substitution possibilities do not vary across many inputs using inflexible forms, the multi-output version of technology using inflexible functions, in general, does not have varying substitution possibilities either. We need more flexible functional representations of production to satisfy the regularity conditions and to resolve the issues of using inflexible forms in the multi-output production. Flexible functional forms allow non-increasing marginal rates of substitution, which is a property all well-defined production functions possess.

One important motivation for using flexible functional forms is that they do not impose any prior restrictions on the Allen-Uzawa elasticities of substitution. Given any arbitrary function, the flexible forms can approximate the function as well as the first two derivatives at a point with precision (Diewert, 1971; Wales, 1977; Fuss et al., 1978; Caves and Christensen, 1980). The flexible functional forms are not completely new knowledge. In fact, they can be derived by adding second-order terms to a wide range of functions used in the production studies. Therefore, the flexible functional forms can be considered as non-parametric versions of the commonly used functional forms such as the linear, the Leontief, and the Cobb-Douglas functions.

In the next section, we will focus on a set of cost functions widely used in the literature: the translog, the quadratic, the generalized Cobb-Douglas, the generalized Leontief, the CES-translog, and the symmetric generalized McFadden cost functions. In this section, we present some of the important features of the dual cost function and issues related to its estimation. Since many of these concepts apply to revenue and profit estimations, we keep those sections relatively brief.

## 3.1 Cost Function Properties

The cost function gives the minimal amount of cost for a certain level of outputs $y \in \mathbb{R}_+^{M_y}$ with given technological possibilities and fixed input prices $w \in \mathbb{R}_{++}^{M_x}$ where $M_y$ and $M_x$ are number of outputs and inputs, respectively. The duality theory shows that the cost function of a productive unit contains all the information of its technology. An immediate example is that the input distance function-based scale elasticity coincides with the cost-based measure of scale elasticity measure. Hence, understanding the cost function is essential for understanding the technology of production. We first summarize the properties of a cost function below as these properties play a central role in estimating a cost function (Varian, 1992; Sickles and Zelenyuk, 2018).

1)  $C(y,w) \geq 0$                  (non-negativity)

2)  $C(y,w)$ is continuous in $(y,w)$   (continuity)[2]

3)  $C(y,kw) = kC(y,w), \forall k > 0$     (linear homogeneity in $w$)

4)  $C(y,\tilde{w}) \geq C(y,w), \forall \tilde{w} \geq w$     (monotonicity in $w$)

5)  $C(y,w)$ is concave in $w$           (concavity in $w$)

where $y \in \mathbb{R}_+^{M_y}$ and $w \in \mathbb{R}_{++}^{M_x}$ are vectors of outputs and input prices, respectively.

In practice, Condition 1 and 2 are automatically satisfied by a proper functional choice for the cost function. Condition 1 may be violated for some functional form choices but, generally, it is satisfied at sample data points. Imposition of Condition 3 is not problematic as well. However, imposing Condition 4 and 5 on a cost function is a relatively more difficult, yet possible, task. The difficulty stems from the fact that, for flexible functional forms, the restrictions would be observation specific. In practice, monotonicity condition is our least concern since estimated factor demands are positive and cost is increasing in output with no parametric restrictions imposed. However, curvature conditions pose a somewhat difficult problem when estimating a flexible functional form.

## 3.2 Functional Forms for Cost Function Estimation

---

[2] A weaker continuity condition is that $C(y,w)$ is continuous in $w$ and lower semi-continuous in $y$.

In this section, we briefly discuss some of the most widely used functional forms for cost function estimation and how regularity conditions are treated in this context. Although we concentrate only on single-output cost functions, the generalizations to multioutput cases are available and straightforward.[3]

### 3.2.1 Translog Cost Function

The translog (TL) cost function (Christensen, et al, 1971) is the most widely used flexible functional form for cost function estimation and is:

$$\ln C(y,w) = \beta_0 + \beta_y \ln y + \sum_j \beta_j \ln w_j + \frac{1}{2}\beta_{yy}(\ln y)^2 + \sum_j \beta_{yj} \ln y \ln w_j + \frac{1}{2}\sum_{j,k}\beta_{jk}\ln w_j \ln w_k \quad (3.2.1)$$

where $\beta_{jk} = \beta_{kj}$ (symmetry), $\sum_j \beta_j = 1$, $\sum_j \beta_{yj} = 0$, and $\sum_k \beta_{jk} = 0$ (linear homogeneity). A standard way to impose linear homogeneity restriction is by normalizing $C(y,w)$ and input prices using one of the input prices. It is common to estimate the cost-input share system in order to add degrees of freedom and boost the precision of the estimates. This, of course, may not be appropriate if input allocations are distorted and thus the cost minimizing input shares derived from the TL are not given by:

$$s_j(y,w) = \frac{\partial \ln C(y,w)}{\partial \ln w_j} = \beta_j + \beta_{yj} \ln y + \sum_k \beta_{jk} \ln w_k. \quad (3.2.2)$$

These input share equations (as opposed to the input demand equations in the level form) are linear in parameters. Regularity conditions can be tested using the cost function's estimates. For example, the monotonicity condition is satisfied if $s_j(y,w) \geq 0$. Linear homogeneity in $y$ is met when $\beta_{yy} = \beta_{yj} = 0$, while the less restrictive property of homotheticity only requires that $\beta_{yj} = 0$. When the TL second-order terms $\beta_{yy}, \beta_{yj}, \beta_{jk}$ are zero we have the Cobb-Douglas (CD) cost function.

### 3.2.2 Translog Cost functions with Allocative and Technical Distortions

---

[3] See Caves et al. (1980) for a discussion multi-output cost functions. See also Röller (1990) for another study that consider multioutput cost functions.

Kumbhakar (1996) discusses inefficiencies with a focus on multiple outputs in the frameworks of cost minimizing and profit maximizing using translog functions to represent technology. A firm minimizes shadow cost given inefficient output and its optimization problem is defined as

$$c^*(w^*, ye^u) = \sum_j w_j^* x_j(w^*, ye^u) \tag{3.2.3}$$

where $w_i^*$ is the shadow price of the optimal input level, $y$ is the actual output, and $u \geq 0$ is technical inefficiency such that $ye^u$ is the maximum possible output. Since shadow costs are unobservable, actual costs are related with shadow costs by using input demand function and are derived as

$$\ln c^A = \ln c^* + \ln\left(\sum_j S_j^* \theta_j^{-1}\right), \tag{3.2.4}$$

where $S_j^*$ is the shadow cost share, and $\theta_j \neq 1$ is allocative inefficiency. Actual cost shares can be related to the shadow cost shares by

$$S_j^A = S_j^* \theta_j^{-1} / \sum_k S_k^* \theta_k^{-1}. \tag{3.2.5}$$

A translog shadow cost function for the shadow cost function is utilized with homogeneity of degree one in $w^*$ and is written as

$$
\begin{aligned}
\ln\left(\frac{c^*}{w_1}\right) = {} & \alpha_0 + \sum_j \alpha_j \ln \widetilde{w}_j^* + \frac{1}{2} \sum_{j,k} a_{jk} \ln \widetilde{w}_j^* \ln \widetilde{w}_k^* \\
& + \sum_m \beta_m \ln(y_m e^u) \\
& + \frac{1}{2} \sum_{m,l} \beta_{ml} \ln(y_m e^u) \ln(y_l e^u) + \sum_{j,m} \gamma_{jm} \ln w_j^* \ln(y_m e^u),
\end{aligned}
\tag{3.2.6}
$$

where $\alpha_{jk} = \alpha_{kj}$, $\beta_{ml} = \beta_{lm}$, and $\widetilde{w}_j^* = \dfrac{w_j^*}{w_1}$.

Then shadow cost shares can be obtained as

$$S_j^* = \partial \ln c^* / \partial \ln w_j^* = \alpha_j + \sum_k \alpha_{jk} \ln \widetilde{w}_k^* + \sum_m \gamma_{jm} \ln(y_m e^u) \tag{3.2.7}$$

Technical inefficiency does not only appear additively but also interact with input prices and outputs, which results in heteroscedasticity.

In the presence of input inefficiency, the shadow cost function incorporating technical inefficiency is

$$\tilde{c}(w^*, y) = \sum_j w_j^* x_j^e(w^*, y), \tag{3.2.8}$$

and input demand functions are derived from Shephard's lemma

$$x_j^e(w^*, y) = \frac{\partial \tilde{c}(w^*, y)}{\partial w_j^*}. \tag{3.2.9}$$

For the translog cost function, actual cost and shadow cost can be related by

$$\ln c^A = \ln \tilde{c}(w^*, y) + \ln(\sum_j \tilde{S}_j \theta_j^{-1}) + \tau, \tag{3.2.10}$$

where $\tilde{S}_j$ is the shadow cost share in the case of input inefficiency. Similar to the derivation in the output inefficiency case, actual cost shares can be derived as

$$S_j^A = \frac{w_j x_j}{c^A} = \tilde{S}_j \theta_j^{-1} / \sum_k \tilde{S}_k \theta_k^{-1}. \tag{3.2.11}$$

The cost function $\ln c^A$ is then complete after using the translog form for $\ln \tilde{c}(w^*, y)$ and $\tilde{S}_j$ is derived from the translog form.

Sickles and Streitwieser (1998) focus on distortions in the pipeline transmission of natural gas by employing a restricted cost function captured by a shadow price and estimate various aspects of a production.

Assuming exogenous output and input prices, a firm minimizes its short run cost as follows

$$\min \sum_i w_i x_i \text{ subject to } G(y, x; t) = 0, \tag{3.2.12}$$

where $G$ is the function that transforms the technology $t$, and $x$ include labor, energy, and two quasi-fixed capital inputs. The solution to this is the short-run variable cost function

$$VC = C(y, w, x; t), \tag{3.2.13}$$

where $C$ is homogenous of degree one, non-decreasing, and concave in factor prices $w$, non-increasing and convex in the quasi-fixed factors $x$, and non-negative and non-decreasing in output $y$. A non-homothetic translog function is used to approximate $C$. Given exogenous $w_i$, they derive the variable cost share utilizing Shephard's Lemma as

$$M_i = \alpha_i + \sum_i \beta_{ij} \ln w_j + \beta_{yi} \ln y + \sum_k \beta_{ik} \ln x_k . \tag{3.2.14}$$

The shadow share equation $-\frac{\partial \ln C}{\partial \ln x_k} = \frac{z_k x_k}{CV}$ is incorporated in the model, where $z_k$, the shadow price, can be obtained by taking the difference between revenues and variable costs. The shadow cost share in the restricted translog cost function is

$$M_k = -[\alpha_k + \sum_i \beta_{ik} \ln w_i + \beta_{yk} \ln y + \sum_h \beta_{hk} \ln x_{hk}]. \tag{3.2.15}$$

Good, Nadiri, and Sickles (1991) develop several modeling scenarios in the airline industry, which allow input price distortions incorporated in a translog variable cost function that captures the linkage between observed cost and assumed minimized cost. Airlines are assumed to use inputs $x = x (x_J, x_{N-J}) > 0$ to produce outputs $y = y (y_K, y_{M-K})$, where the last $N$-$J$ inputs are assumed to be fixed and the last $M$-$K$ outputs are non-physical output characteristics. Consider a virtual technology and virtual input and output decisions, labeled with a '*', that are consistent with the standard assumptions of duality theory. The observed prices deviate from the virtual prices by $\theta = (\theta_1, \dots, \theta_N)$ such that $w_i^* = w_i + \theta_i$ for input $i$. Based on Shephard's lemma, factor demands derived from the firm's minimum virtual cost function are

$$x_J^*(y, w_J^*; x_{N-J}) = \nabla_{w_J^*} C^*(y, w_J^*; x_{N-J}). \tag{3.2.16}$$

The observed cost function and associated short-run factor shares are

$$C(y, w_J^*, w_J; x_{N-J}) = \sum_j w_j x_j^*(y, w_J^*; x_{N-J}) \tag{3.2.17}$$

and

$$M_i = \frac{w_i x_i}{C(y, w_J^*, w_J; x_{N-J})}, \quad i = 1, \dots, J. \tag{3.2.18}$$

Since $M_i^* = w_i^* x_i / C^*$, observed input use can be written as $x_i = M_i^* C^* / w_i^*$. Then, observed costs can be expressed as

$$C = C^* [\Sigma_i \left( \frac{M_i^* w_i}{w_i^*} \right)] , \tag{3.2.19}$$

and observed factor shares expressed as

$$M_i = \frac{\frac{M_i^* w_i}{w_i^*}}{\Sigma_j \left( \frac{M_j^* w_j}{w_j^*} \right)}. \tag{3.2.20}$$

The equations above provide linkages between an observable cost function and the virtual technology when the application of the technology is distorted.

Atkinson and Halvorsen (1984) incorporate regulatory constraints into the cost function framework in which they assume shadow prices to be simply proportional to market prices. Later Getachew and Sickles (2007) utilizes the same approach to study the impact of policy constraints on relative prices and structure of production. By imposing additional constraints $R(w, x; \phi)$, the firm minimizes the production cost as follows

$$\min_x C = w'x \ s.t. \ f(x) \le Q \text{ and } R(w, x; \phi) \le 0$$
$$\tag{3.2.21}$$

where $f(x)$ is a production function, and $Q$ is a certain level of output. Taking Lagrangian, the constrained cost minimization of the firm becomes

$$L = w'x - v(f(x) - Q) - \Sigma_r \lambda_r R_r (w, x; \phi), \tag{3.2.22}$$

where $\lambda_r$ are the Lagrangian multipliers for each of the $R_r$ constraints. The unobserved shadow prices are approximated by using a first-order Taylor series

$$w_i^e = k_i w_i, \tag{3.2.23}$$

where $k_i$ is a factor proportional to an input price. Derived from the shadow cost function, the updated demand function can be obtained utilizing Shepard's Lemma. The updated demand function gives an actual cost function

$$\ln C^A = \ln C^* + \ln \Sigma_i \frac{M_i^*}{k_i}, \tag{3.2.24}$$

where $M_i^*$ is the shadow share of factor $i$. The actual share equation, $M_i^A$, is derived to be

$$M_i^A = \frac{\frac{M_i^*}{k_i}}{\Sigma_i \frac{M_i^*}{k_i}}. \tag{3.2.25}$$

The shadow cost function, $\ln C^*$, can be rewritten in the translog form as follows:

$$\ln C^* =$$

$$\alpha_0 + \alpha_Q \ln Q + \frac{1}{2}\gamma_{QQ}(\ln Q)^2 +$$

$$\Sigma_i \alpha_i \ln(k_i w_i) + \Sigma_i \gamma_{iQ} \ln Q \ln(k_i w_i) + \frac{1}{2}\Sigma_{i,j} \gamma_{ij} \ln(k_i w_i) \ln(k_j w_j) + \delta_t t,$$

(3.2.26)

where $t$ is the time trend that represents technological change over time. Then, the expression for the shadow share $M_i^*$ can be obtained from the logarithmic differentiation. Substituting into the actual cost function gives

$$\ln C^A = \ln C^* + \ln\{\Sigma_i[\alpha_i + \gamma_{iQ} \ln Q + \Sigma_j \gamma_{ij} \ln(k_j w_j)] / k_i\}. \tag{3.2.27}$$

Then, the actual cost share of input $i$ can be derived as

$$M_i^A = [\alpha_i + \gamma_{iQ} \ln Q + \Sigma_j \gamma_{ij} \ln(k_j w_j)]\frac{1}{k_i} / \Sigma_i(\alpha_i + \gamma_{iQ} \ln Q + \Sigma_j \gamma_{ij} \ln(k_j w_j))\frac{1}{k_i}. \tag{3.2.28}$$

The actual cost function is then complete.

### 3.2.3   Generalized Leontief Cost Function

The generalized Leontief (GL) cost function (Diewert, 1971; Diewert and Wales, 1987) is homogenous by construction and is given by:

$$C(y,w) = \sum_j \beta_j w_j + y\sum_{j,k} \beta_{jk} w_j^{1/2} w_k^{1/2} + y^2 \sum_k \beta_{yj} w_j \tag{3.2.29}$$

where $\beta_{jk} = \beta_{kj}$ (symmetry). Input demand equations are given by:

$$x_j(y,w) = \beta_j + y\sum_k \beta_{jk}\left(\frac{w_k}{w_k}\right)^{1/2} + \beta_{yj} y^2. \tag{3.2.30}$$

The monotonicity condition is satisfied if $x_j(y,w) \geq 0$. The GL cost function is non-homothetic unless $\beta_{yy} = \beta_{yj} = 0$ and incapable of distinguishing between homotheticity and linear homogeneity. When $\beta_{jk} = 0$ for $j \neq k$, GL cost function collapses to the Leontief fix proportions cost function.

### 3.2.4 The Symmetric Generalized McFadden Cost Function

The symmetric generalized McFadden (SGM) cost function (Diewert and Wales, 1987) is given by:

$$C(y,w) = g(w)y + \sum_j \beta_j w_j + y \sum_j \beta_{yj} w_j \qquad (3.2.31)$$

where $g(w) = \dfrac{1}{2}\dfrac{w'Sw}{\theta'w}$, $S$ is a symmetric non-negative semidefinite parameter matrix, and $\theta$ is a non-negative vector (not all zero). In order to achieve identification of all parameters, we need to have $S\tilde{w} = 0$ for some $\tilde{w}$ with strictly positive components, e.g., a vector of ones. Input demand equations are given by the vector:

$$x(w) = \frac{Sw}{\theta'w} - \frac{1}{2}\frac{w'Sw}{(\theta'w)^2}\theta. \qquad (3.2.32)$$

By construction, SGM cost function is linear homogenous in $w$. The monotonicity condition is satisfied if the components of $x(w)$ are non-negative. It turns out that $C(y,w)$ is globally concave in $w$ if $S$ is negative semidefinite. If the estimate of $S$ is not negative semidefinite, one can reparametrize $S$ as $S = -LL'$ where $L$ is a lower triangular matrix so that $L'\tilde{w} = 0$, which would assure global concavity of $C(y,w)$. Kumbhakar (1994) gives a generalization of SGM cost function to the multioutput case that makes it relatively easy to estimate different aspects of a production technology. He applies SGM to a panel data of 12 Finnish foundry plants to estimate technical progress, economies of scale, and economies of scope. Rask (1995) proposes a modified version of SGM to allow fixed factors of production so that the cost function can be applied to the processes when there are fixed costs. He estimates the modified SGM cost

function for sugarcane in Brazil, which takes up over two-thirds of total costs in ethanol production and thus is important to study the technology of sugarcane production.

### 3.2.5 Imposing Regularity Conditions for Cost Functions

As Barrett (2002) points out that if both monotonicity and curvature conditions are not satisfied, the second-order conditions for optimization and duality theory fail. While some empirical researchers do not state these conditions, many others are careful about the regularity conditions. Guilkey and Lovell (1980) and Guilkey et al. (1983) exemplify some studies that provide evidence for potential poor global behavior of multioutput cost functions.[4]

If the percentage of violations for monotonicity and curvature conditions is small (e.g., smaller than 5%), some researchers attribute this to the stochastic nature of the estimations and find the violations acceptable. When the percentage of violations is high, some researchers modify the model to get an acceptable violation percentage. For example, when estimating a TL variable cost function of US airports, Kutlu and McCarthy (2016) include an additional term to reduce the violation percentages for monotonicity and concavity conditions. The percentage of violations decrease from 4.2% to 0.5% after including this term. They argue that some airports have particularly higher capital levels relative to the median airport and the additional term that they include captures this pattern.

Another approach is simply imposing regularity conditions. Serletis and Feng (2015) and references there in provide good discussions on how this can be done. Hence, in the rest of this subsection, we closely follow their arguments. Serletis and Feng (2015) categorize these methods as local regularity (at some data point in the sample), regional regularity (over a neighborhood of data points in the sample), pointwise regularity (at every data point in the sample), or global regularity (at all possible data points).

Cholesky decomposition methods for imposing regularity conditions was first used by Wiley et al. (1973). This method is based on the Cholesky decomposition of a Hessian matrix into the product of a lower triangular matrix and its conjugate transpose. For imposing concavity, one can reparametrize a matrix $S$ as $S = -LL'$ where $L$ is a lower triangular matrix. As stated by Serletis and Feng (2015), this approach can be used not only for imposing the curvature but

---

[4] See Wales (1977) for another example in the utility function context.

also the monotonicity conditions. While this approach is capable of imposing local and global curvature conditions, it cannot impose regional or pointwise curvature conditions. For monotonicity, the approach can be used to impose local monotonicity condition. As an illustration, we consider the TL cost function given in Section 3.2.1. The concavity in input prices is satisfied if the Hessian matrix

$$H(y,w) = \frac{\partial^2 C(y,w)}{\partial w \partial w'} \tag{3.2.33}$$

is negative semidefinite. Diewert and Wales (1987) prove that $H$ is negative semidefinite if and only if the following matrix is negative semidefinite:

$$G(y,w) = B - Diag(s(y,w)) + s(y,w)s'(y,w), \tag{3.2.34}$$

where $B = [\beta_{ij}]$ is the matrix with element $ij$ being equal to $\beta_{ij}$, $s(y,w) = \left(s_1(y,w), s_2(y,w), \cdots, s_{M_x}(y,w)\right)'$ is the input share vector, and $Diag(s(y,w))$ is the $M_x \times M_x$ diagonal matrix with diagonal elements being equal to input share vector $s(y,w)$. Since $G(y,w)$ is observation specific, it may not be easy to impose concavity for all data points in the sample. However, as in Ryan and Wales (2000) and Feng and Serletis (2008), we can easily impose concavity on $G(y,w)$ at a reference point in the sample. Usually once the concavity is satisfied at a single reference point, it is satisfied at most of the other sample points (if not all). If the percentage of violations is still high, one can simply try other reference points and find the reference point that gives minimum number of violations. The TL cost function would satisfy global concavity in input prices if $s(y,w) > 0$ and $B$ is negative semidefinite (Diewert and Wales, 1987). However, Lau (1978) and Diewert and Wales (1987) argue that imposing negative semidefiniteness on $B$ destroys the flexibility of TL cost function and reduces it to the Cobb-Douglas form. The imposition of monotonicity by the Cholesky decomposition is not difficult and explained by Serletis and Feng (2015).

The non-linear optimization method for imposing regularity conditions is first used by Geman and Geman (1984). In order to reduce computational difficulties and time, Serletis and

Feng (2015) impose linear homogeneity by normalizing the cost and input prices by the last input price, $w_{M_x}$.[5] They impose negative semidefiniteness on $G(y,w)$, i.e., concavity in input prices, by restricting its eigenvalues to be non-positive. They also impose non-negativity on the cost function and non-negativity of input shares (monotonicity). This approach can impose curvature and monotonicity conditions locally, regionally, and pointwise. It is possible to impose global concavity by restricting the eigenvalues for B to be non-positive. However, the global monotonicity and non-negativity cannot be imposed if we want to keep concavity assumption.

Serletis and Feng (2015) argue that the Bayesian method is a convenient way for imposing regularity conditions due to Gibbs sampling methods introduced by Geman and Geman (1984) and the Metropolis-Hastings algorithm (Metropolis et al., 1953 and Hastings, 1970). Terrell (1996), Koop et al. (1997), and Griffiths et al. (2000) exemplify some important contributions on this area that allow incorporation of non-negativity, monotonicity, and concavity conditions.[6]

Serletis and Feng (2015) examine the performance of all three methods for imposing non-negativity, monotonicity, and concavity conditions for TL cost function. They find that, irrespective of the method, imposing global curvature conditions forces the elements of the *B* matrix to be close to zero as the TL cost function reduces to the Cobb-Douglas cost function in this case. Hence, they rather recommend imposing pointwise regularity using either constraint optimization or Bayesian approach. However, the Bayesian approach may be preferred on the grounds that it is easy to obtain statistical inferences for the parameters and relevant measures (e.g., elasticities and productivity), which can be expressed as functions of parameters.

## 3.3    Stochastic Frontier Models for Cost Functions

The stochastic frontier analysis literature relaxes the neoclassical full efficiency assumption by allowing the productive units to be inefficient. Aigner et al. (1977) and Meeusen and van den Broeck (1977) exemplify earlier studies of stochastic frontier models that aim to measure efficiencies of productive units. A common feature of stochastic frontier models (SFMs) is that they assume a composed error term where the first component is the usual two-sided error

---

[5] For another application of constrained optimization method to a flexible (i.e., globally flexible Fourier) cost function, see Feng and Serletis (2009).
[6] See Kleit and Terrel (2001) as an application of Bayesian approach for flexible cost functions.

and the second component is a one-sided (non-negative) error term, which represents inefficiency. A variety of distributions is proposed for the one-sided error component including the half normal (Aigner et al., 1977), the exponential (Meeusen and van den Broeck, 1977), the truncated normal (Stevenson, 1980), the gamma (Greene, 1980a, 1980b, 2003), and doubly truncated normal (Almanidis et al., 2014) distributions.

A stochastic cost frontier model is given by:

$$\ln C = \alpha + x_1' \beta + u + v \tag{3.3.1}$$

where $C$ is the cost of the productive unit; $\alpha$ is the constant term; $x_1$ is a vector of frontier variables, which does not contain the constant; $u \geq 0$ is the one-sided term that captures the cost inefficiency; $v$ is the usual two-sided error term. It is common to model the inefficiency term as $u = h\left(x_2' \gamma\right) u^*$ where $u^* \geq 0$ is a one-sided random variable and $h > 0$ is a function of so called environmental variables, $x_2$, that affect inefficiency. The smaller values of $u$ indicate that the productive unit is cost efficient, and $u = 0$ means that the productive unit becomes fully efficient. The standard stochastic frontier models assume that $u^*$, $v$, and $\left(x_1', x_2'\right)'$ are all independent from each other. Cost efficiency is estimated by predicting:[7]

$$Eff = \exp\left(-u\right). \tag{3.3.2}$$

The earlier stochastic models (e.g., Aigner et al., 1977 and Meeusen and van den Broeck, 1977) are in the cross-sectional framework. Panel data can potentially give more reliable inefficiency estimates. Pitt and Lee (1981) and Schmidt and Sickles (1984) propose random and fixed effects models for estimating unit specific inefficiencies. These models assume time-invariant inefficiency, which may not be a reasonable assumption for relatively longer panel data. Cornwell et al. (1990), Kumbhakar (1990), Battese and Coelli (1992), and Lee and Schmidt (1993) exemplify earlier time-varying inefficiency models. Ahn et al. (2000), Desli et al. (2003),

---

[7] See Kumbhakar and Lovell (2003) for details.

Tsionas (2006), Huang and Chen (2009), Assaf et al. (2014), and Duygun et al. (2016) provide dynamic efficiency models. Greene (2005a, 2005b) argues that if there is productive unit specific heterogeneity in the frontier and this is controlled, the heterogeneity may be confused with inefficiency. Greene (2005a, 2005b) proposes fixed and random effects models to control for heterogeneity, which are called true fixed effects and true random effects, respectively. The advantage of fixed effects models is that the heterogeneity can be correlated with the regressors. However, it is subject to incidental parameters problem. In particular, while the frontier parameters are consistent, the inefficiency estimates may not be accurate. Wang and Ho (2010) solve this problem by introducing first difference and within transformations to eliminate the fixed effects term. Although the fixed effects models of Greene and Wang and Ho (2005) allow inefficiency to vary over time, the heterogeneity is time-invariant. Kutlu, Tran, and Tsionas (2018) illustrate using Monte Carlo simulations that ignoring time-varying heterogeneity may lead to biased parameter estimates and seriously distorted efficiency estimates. The individual effects model of Kutlu, Tran, and Tsionas (2018) solve this issue by allowing both heterogeneity and inefficiency to vary over time without being subject to incidental parameters problem.

Similar to the conventional cost function estimation, the most widely used functional form in stochastic cost frontier studies is the translog functional form. As stated earlier, in a conventional cost function model if the monotonicity and/or curvature conditions are violated, the second-order conditions for optimization and duality theory fail. The issue is even more serious for stochastic frontier models. Sauer et al. (2006) illustrate that when the monotonicity and curvature conditions are not satisfied, the efficiency estimates may be seriously distorted. Many stochastic frontier studies either do not state whether the regularity conditions are satisfied or simply check these conditions at the mean or median of the sample data points. Hence, the regularity conditions may still be violated at many other sample points, indicating that the cost efficiency estimates for these sample points (and potentially other sample points) are not reliable.

All these stochastic frontier studies mentioned so far can be applied to stochastic cost, production, profit, and revenue frontier model estimations with minor modifications. In particular, for production, profit, and revenue estimations, the inefficiency component $u$ is replaced by $-u$ to estimate efficiency.

Allocative inefficiency results in utilization of inputs in wrong proportions given input prices, i.e., misallocation of inputs. By only using a production function, we can estimate technical efficiency, which happens when the firm fails to produce maximum output from a given input bundle, but we cannot estimate allocative inefficiency. Under the Cobb-Douglass production function assumption, Schmidt and Lovell (1979) present a stochastic cost frontier model where both costs of allocative and technical inefficiency can be estimated. However, they assume that allocative and technical inefficiency are not correlated. Under the same production technology, Schmidt and Lovell (1980) relax this assumption by allowing allocative and technical inefficiencies to be correlated. Modelling allocative inefficiency under translog cost function assumption is less trivial. Greene (1980) models allocative and technical inefficiency in a translog cost function by assuming that allocative inefficiency departures from the cost shares. However, he does not derive cost of allocative inefficiency due to such departures. Rather, he assumes that allocative inefficiency and cost of allocative inefficiency are independent. Bauer (1990) calls this "Greene problem." Kumbhakar and Wang (2006b) and Kutlu (2013) examine the consequences of lumping allocative inefficiency together with technical inefficiency when estimating a cost frontier, i.e., the assumption that the one-sided error term in the cost function captures the overall cost of inefficiency. They both start with the cost minimization problem for the translog cost function. Then, they calculate the exact allocative inefficiency and the corresponding cost of allocative inefficiency where allocative inefficiency is defined as the deviations from the optimal input allocation. Both Kumbhakar and Wang (2006b) and Kutlu (2013) point out negative consequences of lumping the allocative inefficiency with technical efficiency when estimating a cost frontier. Kutlu (2013) argues that system estimators perform worse than single equation estimators even when the complex functional form for allocative inefficiency is approximated by a first order Taylor series. In order to address this issue, Kumbhakar and Tsionas (2005) use similar approximations in a Bayesian setting, and the solutions based on the cost function approach seem not easy. Kumbhakar and Wang (2006a) overcome this issue by using a primal system consisting of a translog production function and first order conditions of cost minimization. In defense of standard stochastic cost frontier models, Kumbhakar and Wang (2006b) and Kutlu (2013) are typical examples for those studies that find negative results based on changing where and how an error term enters a model. While these negative results put some unrest about cost function estimations, they depend on how the data

generating process is determined. Nevertheless, unlike the conventional cost function estimations where researchers generally estimate a cost-input share system, the number of such studies is almost non-existent in the stochastic frontier literature.

## 3.4    Endogeneity in Cost Function Models

Using the production function approach is appropriate if the inputs are exogenous. However, researchers often encounter endogenous input choices in the production process. In particular, the factor inputs under a firm's control may be reallocated to achieve the firm's objectives. In the case of a stochastic production function by a firm maximizing expected profits, (Zellner et al., 1966), all variable inputs can be considered weakly exogenous. However, if the expected profit maximization assumption of Zellner et al. (1966) is not accurate, then one potential solution is to use an instrumental variable or control variable approach to address the issue. In many scenarios, the price taking assumption is more reasonable compared to the exogenous factor inputs assumption and good instruments may be hard to find. Hence, a widely used solution is to estimate a cost function rather than a production function. This is one of the reasons why a dual cost function specification may be preferred over a primal production function specification. Exogenous input prices are more likely when the market is competitive, and thus researchers would prefer the cost function approach given that the level of output is dictated by market forces exogenous to the firm. However, cost functions may suffer from endogeneity problems as well if the output fails to be exogenous. Thus, both production and cost functions may suffer from endogeneity. Besides endogenous outputs, other scenarios may lead to endogeneity in the cost function approach.

One potential problem occurs when a cost function includes a quality variable where the quality is jointly determined by the costs. Mutter et al. (2013) argue that inclusion of the quality variable leads to inconsistent parameter estimates. Some researchers drop the quality variable to avoid such problem, but this does not solve the issue in the stochastic frontier framework. If the quality is cost enhancing and a stochastic frontier model is estimated, the efficiency estimates would be inconsistent irrespective of whether the quality variable is included in the frontier. Duncombe and Yinger (2011) and Gronberg et al. (2011) exemplify studies that point out the endogeneity of output quality in their cost equation. Another potentially endogenous variable

used in cost function estimations is the Herfhindahl-Hirschman Index (HHI). This variable is popular in stochastic frontier models due to close connection between market power and efficiency. In particular, it is common to model inefficiency by using HHI as one of the environmental variables. Karakaplan and Kutlu (2017b, 2018) find evidence of endogeneity from HHI. Similarly, Kutlu, Tran, and Tsionas (2018) find evidence of endogeneity from another related variable that measures profitability, i.e., return on revenue.

The endogeneity problem is more likely to occur in a stochastic frontier setting due to presence of the additional inefficiency term. In particular, as stated earlier, the standard models in this literature assume that $u^*$, $v$, and $\left( x_1^{'}, x_2^{'} \right)^{'}$ are all independent from each other. Guan et al. (2009) and Kutlu (2010) are the earliest studies that aim to solve endogeneity problems in the stochastic frontier setting. These papers relax the independence assumption of $x_1$ and $v$. Guan et al. (2009) achieve this via a two-stage method where in the first stage they get the consistent frontier parameter estimates using the GMM and in the second state they estimate efficiency using a standard stochastic frontier model. Kutlu (2010) uses a limited information maximum likelihood estimation method (single-stage control function estimation) to solve the endogeneity problem. Tran and Tsionas (2013) propose the GMM counterpart of Kutlu (2010). Karakaplan and Kutlu (2017a, 2017b) present cross-sectional and panel data variations of Kutlu (2010) and extend his method to allow environmental variables to be endogenous, i.e., allowing $v$ and $\left( x_1^{'}, x_2^{'} \right)^{'}$ to be correlated. In a Bayesian framework, Griffiths et al. (2016) propose models that allow $v$ and $\left( x_1^{'}, x_2^{'} \right)^{'}$ to be correlated. Using a copula approach, Amsler et al. (2016, 2017) provide cross-sectional models that allow more general correlations, including the correlation between $u^*$ and $\left( x_1^{'}, x_2^{'} \right)^{'}$. The approach requires using a proper copula and may be computationally intensive. Kutlu, Tran, and Tsionas (2018) provide an individual effects panel data model that allows $v$ and $\left( x_1^{'}, x_2^{'} \right)^{'}$ to be correlated, which is a generalization of time-varying heterogeneity as in Wang and Ho (2010). In an appendix, they also provide a copula variation of their model that allows more general correlation structures. However, they argue and illustrate by

Monte Carlo simulations that when the heterogeneity term is included, the consequences of violating general correlation assumptions are not serious if the heterogeneity is controlled.

Finally, the standard modeling of a cost function does not incorporate agency related aspects into the optimization problem. Kutlu, Mamatzakis, and Tsionas (2017) present a model where the manager is a utility maximizer in a quantity-setting oligopoly market. The utility of the manager is a function of profit and her effort level. They assume that higher effort reduces the costs. This introduces an additional structural inefficiency term, which is a specific function of frontier variables. Hence, given that the standard models ignore this structural inefficiency term, the parameter and efficiency estimates from the standard stochastic frontier cost function models would be inconsistent if the assumptions of this model hold. Basically, the solution to this problem would be including the structural inefficiency term as a control function to correct the bias. Gagnepain and Ivaldi (2002, 2017) propose related models where additional terms appear in the cost function due to agency related problems.

## 3.5     Marginal Cost Estimation

Sometimes a researcher is interested in the marginal cost rather than the cost itself. A common application is estimating the cost function and then calculating the marginal cost (e.g., Weiher et al., 2003 and Kutlu and Sickles, 2012). However, in many occasions data on total cost is either not available at all or not available at the desired market level. For example, Weiher et al. (2003), Kutlu and Sickles (2012), and Kutlu and Wang (2018) have airline specific total cost data for the US airlines although these studies are interested in route-airline-specific marginal cost estimates. The new empirical industrial organization literature allows estimation of marginal cost without using total cost data. The marginal cost estimates (along with market power estimates) are obtained by estimating so called conduct parameter (conjectural variations) model where a general form of demand-supply system is estimated. Bresnahan (1989) and Perloff et al. (2007) provide excellent surveys on this topic. Recently, Kutlu and Wang (2018) present a methodology that combines the conduct parameter and stochastic frontier methods that enables estimation of market power, marginal cost, and marginal cost efficiency estimates from a demand-supply system. The advantage of studying marginal cost efficiency over cost efficiency

is that marginal cost efficiency is directly related to deadweight loss. While both measures are valuable, marginal cost efficiency measure may be more relevant from the antitrust point of view.

## 4.    Revenue Functions

In this section, we present some important features of a revenue function and issues related to its estimation. As we mentioned before, since many of the concepts that we introduced apply to the revenue function estimation, this section will be brief.

### 4.1.    Revenue Function Properties

The revenue function gives the maximal amount of revenue a firm can achieve at a certain level of inputs $x \in \mathbb{R}_+^{M_x}$, given technological possibilities and fixed output prices $p \in \mathbb{R}_+^{M_y}$. First, we summarize the properties of a revenue function below as these properties play a central role when we estimate a revenue function (Sickles and Zelenyuk, 2018):

1) $R(x, p) \geq 0$                    (non-negativity)

2) $R(x, p)$ is continuous in $(x, p)$    (continuity)[8]

3) $R(x, kp) = kR(x, p), \forall k > 0$     (linear homogeneity in $p$)

4) $R(x, \tilde{p}) \geq R(x, p), \forall \tilde{p} \geq p$      (monotonicity in $p$)

5) $R(x, p)$ is convex in $p$           (convexity in $p$)

where $x \in \mathbb{R}_+^{M_x}$ *and* $p \in \mathbb{R}_{++}^{M_y}$ are vectors of inputs and output prices, respectively.

In practice, Condition 1 and 2 are automatically satisfied by a proper functional choice for the revenue function. As in the cost function case, Condition 1 may be violated for some functional form choices but, generally, it is satisfied at sample data points. Imposition of Condition 3 is not problematic as well. As in the cost function case, the monotonicity conditions are not problematic in practice. However, again, curvature conditions pose some difficulties when estimating a flexible functional form.

### 4.2.    Functional Forms for Revenue Function Estimation

---

[8] A weaker continuity condition is that $R(x, p)$ is continuous in $p$ and upper semi-continuous in $x$.

Typically, the functional forms used in revenue function estimation are similar to those used in (multiple-output) cost function estimation. Hence, we will be brief in this section. The most widely used revenue function is translog revenue function (Diewert, 1974a), which is given by:

$$\ln R(x, p) = \beta_0 + \sum_j \beta_{xj} \ln x_j + \sum_j \beta_j \ln p_j + \frac{1}{2} \sum_{j,k} \beta_{xxjk} \ln x_j \ln x_k$$

$$+ \sum_{j,k} \beta_{xjk} \ln p_j \ln x_k + \frac{1}{2} \sum_{j,k} \beta_{jk} \ln p_j \ln p_k$$

(4.2.1)

where $\beta_{jk} = \beta_{kj}$, $\beta_{xxjk} = \beta_{xxkj}$ (symmetry), $\sum_j \beta_j = 1$, $\sum_j \beta_{xjk} = 0$, and $\sum_k \beta_{jk} = 0$ (linear homogeneity). The output share equations are given by:

$$s_j^y(x, p) = \beta_j + \sum_k \beta_{xjk} \ln x_k + \sum_k \beta_{jk} \ln p_k.$$

(4.2.2)

Diewert (1974a) provides the details about Generalized Leontief revenue function. A functional form, which we haven't mentioned earlier, that is used in the revenue framework is the mean of order of two revenue functions (Diewert, 1974b). Diewert considers only one input case though the functional form can be extended to a multi-input scenario in a straightforward way. Using solutions to a set of functional equations, Chambers et al. (2013) show that the translog revenue function can be obtained from the Shephard distance function for generalized quadratic functions in the dual price space.

## 4.3. Stochastic Frontier Models for Revenue Functions

Unlike the cost function, the relevant stochastic revenue frontier model needs to be slightly modified and is given by:

$$\ln R = \alpha + x_1' \beta - u + v$$

(4.3.1)

where $R$ is the revenue of a productive unit; $\alpha$ is the constant term; $x_1$ is a vector of input variables; $u \geq 0$ is the one-sided term that captures cost inefficiency; $v$ is the usual two-sided error term. As in the cost function case, the smaller values of $u$ indicate that the productive unit is more cost efficient, and when $u = 0$ the productive unit becomes fully efficient. The standard stochastic frontier assumptions about independence of variables remain the same so that $u^*$, $v$,

and $\left(x_1^{'}, x_2^{'}\right)^{'}$ are all independent from each other. In the case of endogenous input variables, estimates from the revenue function would be inconsistent. The endogeneity solutions mentioned for the stochastic cost frontier models can also be applied to the stochastic revenue frontier function estimation.

Applications of the revenue function are not as prevalent as the cost and production function, but the revenue function is still applicable in various research questions. Kumbhakar and Lai (2016) apply the revenue function to a non-radial and output-specific measure of technical efficiency they propose in a revenue-maximizing framework. They use the maximum likelihood estimation method to estimate a translog revenue-share system. The empirical work by Oliveira and his colleagues (2013) use a revenue function to analyze efficiency of hotel companies in Portugal based on the stochastic frontier approach. Mairesse and Jaumandreu (2005) study the discrepancies between the cross-sectional and time-series estimates of scales and capital elasticities by estimating the production function as well as the revenue function with two panel datasets. They find that the estimates of the functions have little difference and conclude that the bias from other sources, rather than the lack of firm data on output prices, are more likely to be problematic. Rogers (1998) estimate revenue efficiency along with cost and profit efficiency to show the importance of including nontraditional output in bank studies. They find that the standard model understates bank efficiency if nontraditional output is excluded.

## 5.  Profit Functions

In this section, we present some important features of a profit function and issues related to its estimation. We also talk about a less well-known form of profit function, which has many desirable properties, so called alternative profit function.

### 5.1.  Profit Function Properties

The profit function gives the maximal amount of profit for given input and output prices with given technological possibilities. First, we summarize the properties of a profit function below as these properties play a central role when we estimate a profit function (Sickles and Zelenyuk, 2018):

1)  $\pi\left(w, p\right) \geq 0$                       (non-negativity)

2) $\pi(w, p)$ is continuous in $(w, p)$   (continuity)

3) $\pi(kw, kp) = k\pi(w, p), \forall k > 0$     (linear homogeneity in $(w, p)$)

4) $\pi(w, \tilde{p}) \geq \pi(w, p), \forall \tilde{p} \geq p$     (monotonicity in $p$)

5) $\pi(\tilde{w}, p) \geq \pi(w, p), \forall \tilde{w} \leq w$     (monotonicity in $w$)

6) $\pi(w, p)$ is convex in $w$     (convexity in $w$)

7) $\pi(w, p)$ is convex in $p$     (convexity in $p$)

where $w \in \mathbb{R}_{++}^{M_x}$ *and* $p \in \mathbb{R}_{++}^{M_y}$ are vectors of input and output prices, respectively.

While Conditions 2-5 are relatively easily satisfied, the curvature conditions (Condition 6 and 7) and Condition 1 need some extra care. In the banking industry, for example, data points with negative profits are not uncommon. However, profit cannot be negative given a concave production function. To use this result, the profit has to be defined as $\pi(w, p) = p^{'}y - w^{'}x$ and used in the model, instead of reported profit. Observed negative profits violate the property and are problematic.

## 5.2. Functional Forms for Profit Function Estimation

As we discussed earlier in the cost function setting, apparent proper candidates for a profit function are twice differentiable functional forms that are based on a quadratic form. Diewert (1974a) notes that having a second order approximation which is homogenous of degree one is a preferred method. However, in this case, the second order approximation reduces to a first order approximation. Due to this reason, he considers alternatives such as generalized quadratic in square roots profit function and its special case, the generalized Leontief profit function. The extended profit function of Behrman et al. (1992) exemplifies another study that is motivated by the same problem.

Now, we briefly discuss the extended profit function of Behrman et al. (1992). We present this model using their notation. Let $x$ be the vector of variable inputs and $H$ be the quasi-fixed input used for producing multiple output represented by $y$ with prices $p$. We further

combine the output and input prices and quantities as $q = \left(p', w'\right)'$ and $u = \left(-y', x'\right)'$. Then, the generalized Leontief variable profit function can be written as follows:

$$\pi\left(q, H\right) = \sum_{j,k} \gamma_{jk} q_j^{1/2} q_k^{1/2} + \sum_j \gamma_{jH} q_j H^{1/2} \tag{5.2.1}$$

where $\gamma_{jk} = \gamma_{kj}$ . Therefore, the constant elasticity transformation-constant elasticity of substitution- generalized Leontief variable profit function (CET-CES-GL) can be expressed as:

$$\pi\left(q, H\right) = \left(\sum_j \gamma_{jj} q_j^{\varepsilon}\right)^{1/\varepsilon} + \sum_{j,k \neq j} \gamma_{jk} q_j^{1/2} q_k^{1/2} + \sum_j \gamma_{jH} q_j H^{1/2} \tag{5.2.2}$$

where $\gamma_{jk} = \gamma_{kj}$. The corresponding variable profit maximizing output supply and input demand equations are given by:

$$u_j\left(q, H\right) = \gamma_{jj} q_j^{\varepsilon-1} \left(\sum_k \gamma_{kk} q_k^{\varepsilon}\right)^{(1-\varepsilon)/\varepsilon} + \sum_{k \neq j} \gamma_{jk} q_j^{-1/2} q_k^{1/2} + q_{jH} H^{1/2}. \tag{5.2.3}$$

### 5.3. Profit Function with Allocative and Technical Distortions

Lovell and Sickles (1983) incorporate technical and allocative inefficiency into a profit function in the Generalized Leontief form by assuming wrong price ratios and by allowing the actual output and input to differ from the optimal levels. The output prices $p = (p_1, ..., p_m) > 0$ and input prices $w = (w_1, ..., w_n) > 0$ are given as exogenous, the profit maximization problem becomes

$\max_{y,x} py - wx$ s.t. $(y, -x) \in T$.

(5.3.1)

The profit function is useful from the fact that a profit function $\pi$ and a production possibilities set $T$ both represent the profit-maximizing technology due to a duality relationship. In addition, profit maximizing output and input allocations can be derived using Hotelling's Lemma:

$\nabla_p \pi(p, w) = y(p, w),\ \ \nabla_w \pi(p, w) = -x(p, w).$ \hfill (5.3.2)

The profit of a firm producing two outputs using two inputs, as an example, is assumed to be the Generalized Leontief form. Then, the profit maximizing output and input equations can be derived from Hotelling's Lemma and can be modified to include inefficiency as follows

$$y_1(p,w,\phi,\theta) = (\beta_{11} - \phi_1) + \beta_{12}\left(\theta_{12}\frac{p_1}{p_2}\right)^{-\frac{1}{2}} + \beta_{13}\left(\theta_{13}\frac{p_1}{w_1}\right)^{-\frac{1}{2}} + \beta_{14}\left(\theta_{14}\frac{p_1}{w_2}\right)^{-\frac{1}{2}}, \tag{5.3.3}$$

$$y_2(p,w,\phi,\theta) = (\beta_{22} - \phi_2) + \beta_{12}\left(\theta_{12}\frac{p_1}{p_2}\right)^{\frac{1}{2}} + \beta_{23}\left(\theta_{23}\frac{p_2}{w_1}\right)^{-\frac{1}{2}} + \beta_{24}\left(\theta_{24}\frac{p_2}{w_2}\right)^{-\frac{1}{2}}, \tag{5.3.4}$$

$$-x_1(p,w,\phi,\theta) = (\beta_{33} - \phi_3) + \beta_{13}\left(\theta_{13}\frac{p_1}{w_1}\right)^{\frac{1}{2}} + \beta_{23}\left(\theta_{23}\frac{p_2}{w_1}\right)^{\frac{1}{2}} + \beta_{34}\left(\theta_{34}\frac{w_1}{w_2}\right)^{-\frac{1}{2}}, \tag{5.3.5}$$

$$-x_2(p,w,\phi,\theta) = (\beta_{44} - \phi_4) + \beta_{14}\left(\theta_{14}\frac{p_1}{w_2}\right)^{\frac{1}{2}} + \beta_{24}\left(\theta_{24}\frac{p_2}{w_2}\right)^{\frac{1}{2}} + \beta_{34}\left(\theta_{34}\frac{w_1}{w_2}\right)^{\frac{1}{2}}. \tag{5.3.6}$$

The parameters $\phi_i \geq 0$ measure the under-production of outputs and excessive usage of inputs due to technical inefficiency. The parameters $\theta_{ij} > 0, j > i$ are interpreted as allocative inefficiency. If both technical and allocative inefficiency exist, the observed profit can be expressed

$$\pi(q,\phi,\theta) = \sum_i^4 (\beta_{ii} - \phi_i)q_i + \sum_i^3\sum_j^4 \beta_{ij}\left(\theta_{ij}^{-\frac{1}{2}} + \theta_{ij}^{\frac{1}{2}}\right)q_i^{\frac{1}{2}}q_j^{\frac{1}{2}}, \tag{5.3.7}$$

where $q \equiv (p_1, p_2, w_1, w_2)$. The change in profit due to technical inefficiency is obtained by

$$\pi(q) - \pi(q,\phi) = \sum_i^4 \phi_i q_i, \tag{5.3.8}$$

and the change in profit due to allocative inefficiency is obtained by

$$\pi(q) - \pi(q,\theta) = \sum_i^3\sum_j^4 \beta_{ij}\left(2 - \left(\theta_{ij}^{-\frac{1}{2}} + \theta_{ij}^{\frac{1}{2}}\right)\right)q_i^{\frac{1}{2}}q_j^{\frac{1}{2}}. \tag{5.3.9}$$

Allocative inefficiency can be further decomposed into output mix inefficiency, input mix inefficiency, and scale inefficiency depending on $\theta_{ij}$. The perceived price ratios, $(\theta_{ij}\frac{q_i}{q_j})$, are consistent allocative inefficiency if they satisfy

$$\left(\theta_{ij}\frac{q_i}{q_j}\right)\left(\theta_{jk}\frac{q_j}{q_k}\right) = \theta_{ik}\frac{q_i}{q_k}, \quad i < j < k, \tag{5.3.10}$$

which requires

$$\theta_{ik} = \theta_{ij}\,\theta_{jk}, \quad i < j < k. \tag{5.3.11}$$

Based on the work of Lovell and Sickles (1983), Sickles, Good, and Johnson (1986) apply the Generalized Leontief profit function with allocative distortions to the US airline industry by assuming wrong price ratios. The generalized Leontief profit function including output characteristics is expressed as

$$\pi(q,c,t;\theta) =$$

$$\sum_i \beta_{ii}q_i + \sum_{i,j}\beta_{ij}\left(\theta_{ij}^{-\frac{1}{2}} + \theta_{ij}^{\frac{1}{2}}\right)q_i^{\frac{1}{2}}q_j^{\frac{1}{2}} + \sum_i \beta_{it}q_i t + \sum_{i,j,k}\delta_{ijk}q_i c_j^{\frac{1}{2}}c_k^{\frac{1}{2}}, \qquad \delta_{ijk} = \delta_{ikj}, \forall i,j \neq$$

$$k, \qquad\qquad\qquad\qquad\qquad\qquad (5.3.12)$$

where $q$ is the vector of input and output prices, $c$ is the vector of output characteristics, and $t$ is a time index. The output and input allocation equations can be derived as

$$d_i(q,c,t;\theta) = \beta_{ii} + \sum_{j\neq i}\beta_{ij}\left(\theta_{ij}\frac{q_i}{q_j}\right)^{\frac{1}{2}} + \beta_{it}t + \sum_{j,k}\delta_{ijk}c_j^{\frac{1}{2}}c_k^{\frac{1}{2}}, \qquad\qquad (5.3.13)$$

where $d = (y, -x)$. The output characteristics are approximated by

$$c_i(q,t) = \sum_j \sum_{k>j}\gamma_{ijk}q_j^{1/2}q_k^{-\frac{1}{2}} + \sum_j \sum_{k>j}\gamma_{ijkt}q_j^{\frac{1}{2}}q_k^{-\frac{1}{2}}t + \gamma_{it}t + \gamma_i. \qquad (5.3.14)$$

Kumbhakar (1996) models technical and allocative inefficiencies in profit maximizing frameworks emphasizing on multi-outputs and multi-inputs. He derives the exact relations between the inefficiencies and profit when translog functions are used to represent technology. In the presence of output technical inefficiency, the firm's profit maximization problem is

$$\max_{y,x} \pi = p'y - w'x$$

$$\text{s.t. } F(ye^u, x) = 0, \qquad\qquad\qquad\qquad\qquad\qquad (5.3.15)$$

where $y$ is the actual output, and $u \geq 0$ is the technical inefficiency so that $ye^u$ is the maximum possible output level. Assume $w_j^* = \theta_j w_j$ and $p_m^* = k_m p_m$, where $\theta_j$ and $k_m$ are input inefficiency and output inefficiency respectively. Optimal inputs and outputs are determined by the shadow profit adjusted for efficiency. The efficiency adjusted normalized shadow profit is

$$\hat{\pi}^* = y_1 e^u + \sum_m \tilde{p}_m^* y_m e^u - \sum_j \hat{w}_j^* x_j = \hat{\pi}^*(\hat{w}^*, p^*) \qquad\qquad (5.3.16)$$

where $\hat{\pi}^* = \frac{\pi^* e^u}{p_1^*}$, $\hat{w}_j^* = e^u \tilde{w}_j^* = \frac{e^u w_j^*}{p_1}$, $\tilde{p}_m^* = p_m^*/p_1$ and $p_1^* = p_1$. The normalized actual profit adjusted for efficiency and the shadow profit adjusted for efficiency are related as follows

$$e^u \tilde{\pi}^A = \hat{\pi}^*[1 + \sum_m \left(\frac{1}{k_m} - 1\right)R_m^* + \sum_j \left(\frac{1}{\theta_j} - 1\right)Q_j^*], \qquad (5.3.17)$$

where the shadow revenue and cost shares are $R_m^* = \frac{\partial \ln \tilde{\pi}^*}{\partial \ln \tilde{p}_{mj}^{**}}$, $Q_j^* = \frac{\partial \ln \hat{\pi}^*}{\partial \ln \hat{w}_j^*} = -\frac{\tilde{w}_j^* x_j}{\tilde{\pi}^*}$. This transforms into

$$\ln \tilde{\pi}^A = \ln \hat{\pi}^* + \ln H - u, \qquad\qquad\qquad\qquad (5.3.18)$$

where $H$ incorporates the shadow revenue and cost shares. The equations that relate the actual revenue and cost shares to the shadow revenue and cost shares are given by

$$R_m^A = R_m^* \frac{1}{H} \frac{1}{k_m} \tag{5.3.19}$$

$$Q_j^A = -Q_j^* \frac{1}{H} \frac{1}{\theta_j}. \tag{5.3.20}$$

Using a translog form for $\hat{\pi}^*(\hat{w}^*, p^*)$ gives the expressions for the shadow revenue and cost shares, we can obtain the expression for $H$. The profit function specification is then complete.

In the presence of input technical inefficiency, the firm maximizes the profit as follows

$$\max_{y,x} \pi = p'y - w'x$$

$$s.t. F(y, xe^{-\tau}) = 0, \tag{5.3.21}$$

where $\tau \geq 0$ is interpreted as technical inefficiency and $e^{-\tau} \leq 1$ as input technical efficiency. Similar to the output technical inefficiency setup, the normalized shadow profit function is

$$\tilde{\pi}^*(w^* e^\tau, p^*) = \frac{\pi^*(.)}{p_1} = y_1 + \sum_m \tilde{p}_m^* y_m - \sum_j \hat{w}_j^* x_j^e, \tag{5.3.22}$$

where $w_j^* = \frac{w_j^* e^\tau}{p_1}$ and $x_j^e = x_j e^{-\tau}$. Since $\tilde{\pi}^*(w^* e^\tau, p^*)$ is not observed, it can be related to the normalized actual profit by

$$\tilde{\pi}^A = \tilde{\pi}^* (1 + \sum_m \left(\frac{1}{k_m} - 1\right) \tilde{R}_m + \sum_j \left(\frac{1}{\theta_j} - 1\right) \tilde{Q}_j), \tag{5.3.23}$$

which implies $\ln \tilde{\pi}^A = \ln \tilde{\pi}^* + \ln \tilde{H}$ where $\tilde{p}_m = \frac{p_m}{p_1}$, $\tilde{w}_j = \frac{w_j}{p_1}$, $\hat{w}_j = w_j e^\tau$. Same procedure follows as in the output technical inefficiency case in which the derived shadow revenue and cost shares can be related to the actual shares. Assuming a translog form for $\tilde{\pi}^*$ gives expressions for the shadow revenue and cost shares.

## 5.4.    Stochastic Frontier Models for Profit Functions

The stochastic frontier models for profit functions differ from the models for cost and revenue functions in the presence of technical inefficiency. Kumbhakar (2001) derive the expressions for the profit function corresponding to different assumptions on the underlying production function. In the presence of technical inefficiency, the profit function can be written as $\pi (p, w, u) = \pi (w, pe^{-u})$, where p is the output price, w is the input price, and $e^{-u} \leq 1$ is a

measure of technical inefficiency. To illustrate, we assume a translog form on actual profit. We write the estimable profit function as follows:

$$\ln\left(\frac{\pi}{p}\right) = \alpha + \Sigma\alpha_j \ln\left(\frac{w_j}{pe^{-u}}\right) + \frac{1}{2}\Sigma\Sigma\alpha_{jk} \ln\left(\frac{w_j}{pe^{-u}}\right)\ln\left(\frac{w_k}{pe^{-u}}\right) - u + v \qquad (5.4.1)$$

or in terms of the profit frontier:

$$\ln\left(\frac{\pi}{p}\right) = \ln\pi(p,w) + \ln h(p,w,u) + v \text{ , where} \qquad (5.4.2)$$

$$\ln h(p,w,u) = -u\{1 - \Sigma\alpha_j - \Sigma\Sigma\alpha_{jk}\ln\left(\frac{w_j}{p}\right) - \frac{u}{2}\Sigma\Sigma\alpha_{jk}\} \qquad (5.4.3)$$

is profit technical inefficiency, which is not a constant multiple of u unless $\Sigma\alpha_{jk} = 0 \ \forall k$, i.e. the underlying production technology is homogenous.

The standard stochastic profit frontier models assume that u, v, and the profit frontier variables are all independent from each other. These assumptions can be relaxed as stated in the stochastic cost frontier section.

In empirical applications, negative accounting profit is a commonly observed phenomenon. However, the dependent variable for a stochastic profit frontier model is the logarithm of the profit, which is not defined for observations with negative profit. Some studies drop the observations with negative profits and estimate the model with the remaining observations. As Bos and Koetter (2009) mention, this method has at least two shortcomings. First, we cannot obtain efficiency estimates for the observations that we drop. Second, these observations are likely to belong to the least efficient productive units. Hence, dropping these observations may potentially distort efficiency estimates. An alternative method is rescaling $\pi$ for all firms so that the rescaled $\pi$ becomes positive. For example, a commonly used recalling is done by adding $\theta = \min\left(\left\|\pi^-\right\|\right) + 1$ to $\pi$ where $\pi^- = \min\left(\pi, 0\right)$ is the negative part of $\pi$. Hence, the stochastic frontier profit model is given as follows:

$$\ln\left(\pi + \theta\right) = \alpha + x_1'\beta - u + v. \qquad (5.4.4)$$

Berger and Mester (1997), Vander Vennet (2002), Maudos et al. (2002), and Kasman and Yildirim (2006) exemplify some studies that use this rescaling approach. Critics to this approach

would ask "Where did this money come from?" Hence, Berger and Mester (1997) modify the prediction of profit efficiency as follows:

$$Eff_{\pi} = \frac{\hat{f}_{\pi} - \widehat{u} - \theta}{\hat{f}_{\pi} - \theta} \tag{5.4.5}$$

where $\hat{f}_{\pi} - \widehat{u} - \theta$ is the predicted actual profit and $\hat{f}_{\pi} - \theta$ is the predicted maximum of profit that could be earned if the productive unit is fully efficient. In order to reflect the actual amounts, the profits are adjusted by $\theta$ and thus the standard formula for efficiency calculations does not work.[9]

Finally, an issue in the estimation of a stochastic frontier profit function is that the risk needs to be included in the model when the production involves risks (Hughes and Mester, 1993, Hughes et al., 1995, and Clark, 1996). Since the risk-taking behavior of a productive unit represents its objective, we would incorrectly consider the risk-averse productive units as relatively inefficient when the risk is not included in the estimation. The studies on financial sectors (e.g., banking) are generally careful about controlling for risk when estimating a profit or alternative profit function.

## 5.5.    Alternative Profit Function

Alternative profit function, introduced by Humphrey and Pulley (1997), is another representation of profits that can be used when the underlying assumptions of standard profit function do not hold. In contrast to the profit function, which takes input and output prices as given, the alternative profit function takes the input prices and output as given, i.e., $\pi(w, y)$. Hence, the independent variables for an alternative profit function are the same as that of a cost function. The underlying assumption in derivation of the alternative profit function is that the

---

[9] Bos and Koetter (2009) propose an alternative approach to overcome this issue. For observations where the profit is positive, they keep the left-hand-side variable as $\ln \pi$ and for those observations where the profit is negative, they replace the left-hand-side variable with 0. They also add an indicator variable to the right-hand-side. This indicator variable equals 0 when the profit is positive and equals $\ln|\pi^{-}|$ when the profit is negative. This method has the advantage that it uses all sample points for the estimations. However, when measuring inefficiency, the logarithmic scale breaks down for negative profits. Hence, the interpretation of inefficiency estimates for the observations with negative profits deviates from the standard interpretation. Koetter et al. (2012) paper exemplifies a study that uses this approach.

productive units maximize profits by choosing input quantities and output prices. Berger and Mester (1997) list four conditions where estimating alternative profit function may provide useful information:

    (1) There are substantial unmeasured differences in quality of outputs.

    (2) Outputs are not completely variable so that the productive unit cannot achieve every scale and output mix.

    (3) Output markets are not perfectly competitive.

    (4) Output prices are not accurately measured.

A model of alternative profit function is very similar to that of a cost function except the dependent variable and the linear homogeneity in input prices assumption. However, in the stochastic frontier setting, an alternative stochastic profit model does not penalize high-quality banks in terms of efficiency, which may not be the case for a stochastic frontier cost model.

It is important to note that, unlike the profit function, the alternative profit function is not linearly homogenous in input prices (Restrepo-Tobón and Kumbhakar, 2014). Hence, linear homogeneity of an alternative profit function is an empirical question and not a theoretical restriction. Restrepo-Tobón and Kumbhakar (2014) illustrate that incorrect imposition of linear homogeneity in prices may lead to misleading results.

## 6.      Multi-output Functional Forms

In productivity analysis, we need data on input and output levels to estimate the production function. The difficulty in obtaining input data and the fact that more companies have integrated production across different segments make it even harder to access to division-level input information. Data on total output and input do not show how the company allocates resources in a certain segment, and thus we cannot estimate the production function for one specific segment. Same problem persists when we study a country's productivity. In this case, most models of productivity assume one common production function for the whole economy. This does not correctly reflect how a country invests its resources since different industries/ sectors use technology differently.

Gong and Sickles (2017) develop modeling and estimation methods for multidivisional/multiproduct firms and improve standard assumptions in the productivity and efficiency literature. They develop a model to find input allocations among different divisions

given total inputs, outputs from each division, and input prices averaged over the segment. The stochastic frontier model for a company $i$ at time $t$ is

$$y_{it} = f(x_{it}; \beta_0)e^{z\tau}e^{v_{it}}e^{-u_{it}} \tag{6.1}$$

where $y_{it}$ is the total output (aggregated); $x_{it} = (x_{it}^1, x_{it}^2, \ldots, x_{it}^M)$ is the vector of inputs of $M$ types; $f(x_{it}; \beta_0)e^{z\tau}$ is the average production frontier, $\beta_0 = (\beta_{01}, \beta_{02}, \ldots, \beta_{0M})$ is a vector of $M$ types of parameters, $z$ is a vector including time dummy variables, and $\tau$ is a vector of corresponding coefficients; $e^{v_{it}}$ is the random shocks to the production, and $u_{it}$ is a one-sided stochastic term related to technical efficiency.

To allow different frontiers for different segments, Gong and Sickles (2017) introduce segment-specific production frontier (SSPF). In an economy that produces $N$ outputs/segments using $M$ inputs during $T$ periods, the production technology in each segment is characterized in a system of $N$ equations as follows:

$$\begin{cases} y_{i1t} = f_1(x_{i1t}; \beta_1)e^{z_1\tau_1}e^{v_{i1t}}e^{-u_{i1t}} \\ \quad \vdots \\ y_{iNt} = f_N(x_{iNt}; \beta_N)e^{z_N\tau_N}e^{v_{iNt}}e^{-u_{iNt}} \end{cases} \tag{6.2}$$

where $y_{ijt}$ is the observed output and $x_{ijt}$ are vectors of inputs (unobserved) of firm $i$ in segment $j$ at time $t$. The production frontier for segment $j$ is represented by $f_j(x_{ijt}; \beta_j)$. Note that the parameters in the production, $\beta_j$, are segment specific. Similar to the single frontier case, $z_j = (z_{j2}, z_{j3}, \ldots, z_{jT})$ is a vector of year dummy variables, and $\tau_j = (\tau_{j2}, \tau_{j3}, \ldots, \tau_{jT})$ is a vector of the corresponding coefficients. The technical efficiency, $u_{ijt} = -\eta(t - T)u_{ij}$, is time variant. The random shock, $v_{ijt}$, is assumed to be drawn from $N(0, \sigma_{vj}^2)$. We can use the SSPF framework to predict division-level efficiency, $\widehat{TE}_{ijt} = e^{-u_{ijt}}$. In the case of single-frontier, the SSPF approach predicts firm-level efficiency. It is straightforward to see that the firm-level efficiency for the multidivisional firm, $\widehat{TE}_{it}$, is the average of division-level efficiency weighted by the ratio of division-level revenue to firm-level revenue:

$$\widehat{TE}_{it} = \sum_j \left[ \frac{R_{ijt}}{R_{it}} \widehat{TE}_{ijt} \right]. \tag{6.3}$$

The SSPF approach incorporates the heterogeneity in production frontiers and has advantage in deriving division-level efficiency compared with a traditional SPF.

## 7.    Non-parametric Estimation (and Shape Restrictions):

No better example exists of a disconnect between the conditions under which a dual relationship is estimated and interpreted than in the case of non-parametric estimation of cost functions. Such relationships have been estimated in the literature for a variety of important industries, most notably in banking services where substantial data in the form of panels of cross sections are publicly available.

This issue has been well studied over the last several decades. A number of important papers have contributed to the development of shape restrictions in non-parametric estimation. A short list includes Matzkin (1991, 1994), Ruud (1997), Fox (1998), Mammen and Thomas-Agnan (1999), Hall and Huang (2001), Ait-Sahalia and Duarte (2003), Lewbel (2010), Shively et al. (2011), Du et al. (2013), and Wu and Sickles (2018).

We begin with some examples to show how to restrict a function by transforming the function. First, we look at how we restrict a function's range. If we want a function to be nonnegative, for example, a common approach is to specify the function as $f(x) = (r(x))^2$ or $(x) = e^{r(x)}$ such that $f(x) \geq 0$. To further restrict the values of the function to be (0,1), we can specify the function as $(x) = \frac{1}{1+e^{r(x)}}$ such that $0 < f(x) < 1$. In general, a range restriction on a function to take values in (a,b) is $f(x) = a + \frac{b-a}{1+e^{r(x)}}$ such that $a < f(x) < b$. In this way, we transform a constrained problem (specifying $f$) into an unconstrained problem (specifying $r$) and still maintain global compliance with constraints. This contrasts with the kernel-based methods which keep observation-specific compliance as seen in Mukerjee (1988) and Mammen (1991). Later studies on penalized kernel-smoothers include Hall (2001), Henderson (2012), Blundell et al. (2012), Ma (2013), and Du (2013).

To impose monotonicity constraints, we utilize integration procedures. Suppose we have a monotone function $f(x)$ and $x \in [0,1]$. Ramsay (1998) represents $f(x)$ as:

$$f(x) = \int_0^x e^{r(s)} ds. \tag{7.1}$$

Monotonicity and concavity can be imposed on $f(x)$ by introducing an unconstrained function $r(x)$, such that $f'(x) = e^{r(x)} > 0$ and $f''(x) = f'(x)r'(x)$ so that if $r'(x) < 0$, then $f''(x) < 0$. One way to model $r(x)$ is to use an integration transformation as follows

$$f(x) = \int_0^x \exp\left(\underbrace{-\int_0^s g(t)dt}_{r(s)}\right)ds. \tag{7.2}$$

It is clear that $f'(x) = \exp\left(-\int_0^x g(t)dt\right) > 0$ and $f''(x) = -f'(x)g(x)$. Therefore, if $g(x)>0$ monotonicity and concavity naturally follow. Such a positivity constraint can be imposed via functions such as $g=x^2$ or $g=exp(x)$.

Wu and Sickles (2017) utilize a spline basis for the non-parametric expression for the function $g(x) = g(h(x))$ where, e.g., $g(x) = (h(x))^2$. A d-th order splines can be written as

$$\Gamma(x) = (\underbrace{1, x, \dots, x^d}_{\text{power series}}, \underbrace{(x-j_1)_+^d, \cdots, (x-j_M)_+^d}_{\text{piecewise power series}})^T, \tag{7.3}$$

where $(x)_+ = \max(x, 0)$ and $j_1 < \cdots < j_M$ are spline knots and thus $h(x) = c^T\Gamma(x)$ where $c$ are the spline coefficients. This leads to a non-parametric production model with monotonicity and curvatures constraints given by:

$$\begin{aligned} y_i &= \beta_0 + \beta_1 \int_0^{x_i} \exp\left(-\int_0^s g\left(c^T\Phi(t)\right)dt\right)ds + \varepsilon_i \\ &= f(x_i; \beta, c) + \varepsilon_i. \end{aligned} \tag{7.4}$$

Then, we can derive the penalized nonlinear least squares estimator as:

$$\min_{\beta,c} \underbrace{\frac{1}{n}\sum_{i=1}^n(y_i - f(x_i; \beta, c))^2}_{\text{goodness of fit}} + \underbrace{\lambda R(f)}_{\text{roughness penalty}}, \tag{7.5}$$

where $R(f) > 0$ measures roughness in $f$ and $\lambda$ controls the balance between the goodness-of-fit and smoothness such that we don't over fit the model. The spline coefficients become closer to 0 as $\lambda R(f)$ decreases. We can use the common integrated squared derivatives, $R(f) = \int_0^1 (f^{(q)}(x))^2 dx, q = 1, 2, \dots$ to model the penalty.

Simar, Van Keilegom, and Zelenyuk (2017) propose a non-parametric least-squares method, which utilize the advantage of the local MLE method with less strict assumptions (and

less computational complexity) to analyze efficiency in the context of stochastic frontier. Given a set of i.i.d. random variables $(Y_i, X_i, Z_i), i = 1, \ldots, n$ where $Y_i$ is the output, $X_i$ are the inputs, and $Z_i$ can be considered as environmental conditions that affect the production. Setting $X_i = x$, and $Z_i = z$, we can characterize the output produced as:

$$Y = m(x, z) - u + v, \tag{7.6}$$

where $m(x, z)$ is the production frontier unknown to researchers, $u \sim D^+(\mu_u(x, z), var_u(x, z))$, and $v \sim D(0, var_v(x, z))$ are independent random variables conditional on $(X, Z)$. The method first estimates an average production function along with some moments of the complex error term. Then the local inefficiency can be computed after identifying the local asymmetry of the error. To estimate the average production function in the first step, we rewrite the output equation as:

$$Y = r_1(x, z) + e, \tag{7.7}$$

where $r_1(x, z) = m(x, z) - \mu_u(x, z)$ represents the average production function, $e = v - u + \mu_u(x, z)$. We can estimate $r_1(x, z) = E(Y|x, z)$ using standard non-parametric methods and obtain the residuals $\hat{e}_i = Y_i - \hat{r}_1(X_i, Z_i), i = 1, \ldots, n$. Then, we can consistently estimate $r_2(x, z)$ and $r_3(x, z)$ where $r_j(x, z) = E(e^j | x, z)$ by:

$$\hat{r}_j(x, z) = \sum W_{i,h}(x, z)\, \hat{e}_i^j, \tag{7.8}$$

where $W_{i,h}(x, z)$ represents the estimation method that depends on the vector of bandwidth $h$. Next, we assume semi-parametric forms of independent $u$ and $v$ as:

$$u|x, z \sim N^+(0, \sigma_u^2(x, z)), \tag{7.9}$$

$$v|x, z \sim N(0, \sigma_v^2(x, z)). \tag{7.10}$$

We can derive expressions in terms of $r_2$ and $r_3$ for $\sigma_u^3(x, z)$ and $\sigma_v^2(x, z)$, and thus we obtain the variance estimates $\hat{\sigma}^2{}_u(x, z)$ and $\hat{\sigma}^2{}_v(x, z)$ by plugging $\hat{r}_2$ and $\hat{r}_3$ estimated in the first step. Then, individual efficiency scores can be estimated using the method of Jondrow et al (1982) for the general case. We can also consistently estimate the conditional mean of inefficiency

term $\hat{\mu}_u(x, z)$ after we obtain the variance estimates. Then, the stochastic frontier can be estimated by:

$$\hat{m}(x, z) = \hat{r}_1(x, z) + \hat{\mu}_u(x, z). \tag{7.11}$$

Non-parametric methods have become essential to analyze productivity. Developments in computational software and statistical methods have contributed greatly to empirical studies on productivity. The challenge remains, however, that estimating non-parametric models is more difficult than imposing restrictions needed for interpreting results from different functional forms.

As we pointed out in our introductory remarks, our chapter focuses on "how cost, revenue, and profit functions are used to identify and characterize an underlying technology" and that we "concentrate on the more widely used cost functions to motivate various issues." Thus, we focus on parametric functions. There are of course many alternative nonparametric methods to specify both the mean function and the error terms in a stochastic frontier function that can be utilized to estimate the dual cost, revenue, and profit functions or the primal production or distance function. Relatively recent work has also focused on methods to ensure that the regularity conditions in such nonparametric approaches are imposed on the functions estimated. This literature includes the work by Fan, Li, and Weersink (1996), Adams, Berger, and Sickles (1997, 1999), Adams and Sickles (2007), Kuosmanen and Kortelainen (2012), and Simar, Van Keilegom, and Zelenyuk (2017).

## 8. Concluding Remarks

What we have covered is just a small part of the economic theory and practice enriched by the development of the duality theory. To conclude this chapter, we would like to emphasize some of the benefits from the duality theory of the production function. We use the cost function to summarize the benefits. Revenue and profit functions have similar properties and the benefits can also be applied. First, the cost function enables us to derive an easier representation of technology. In the case of multiple outputs, for example, the production function becomes infeasible and only yields an implicit function. However, we are able to use the cost function, which has become a common practice. Second, the dual approach using the cost function incorporates optimal input allocation from optimizing firms' behavior, while the primal approach which uses the distance or production function does not contain such information. Third, we can

check if a firm is a cost-minimizer based on the precise conditions that the cost function needs to satisfy in order to characterize the technology of a cost-minimizing firm. In addition, we are able to specify a functional form of the cost function suitable for our estimations as long as it satisfies all the precise conditions. Finally, the data on outputs and input prices required to estimate the cost function are easier to obtain than the data on actual output and input levels required in the estimation using the primal approach.

We also have illustrated how to specify functional forms of technology in various optimization problems which are consistent with both primal and dual relationships in empirical productivity studies. One important lesson is that allowing flexible functional forms may sacrifice parsimony properties. Researchers always need to consider the benefits and losses from flexibility when choosing functional forms to represent the production technology. Last but not least, we must be consistent when interpreting results from any productivity research with the standards established forty years ago that are still of great importance today.

## References

Adams, R.M, Berger, A., Sickles, R.C. (1997), Computation and inference in semiparametric efficient estimation, Computational Approaches to Economic Problems, Advances in Computational Economics, edited by Hans Amman, Berc Rustem and Andrew Whinston, Boston: Kluwer Academic, 57-70.

Adams, R.M, Berger, A., Sickles, R.C. (1999), Semiparametric approaches to stochastic panel frontiers with applications to the banking industry, *Journal of Business and Economic Statistics*, 17, 1999, 349-358.

Adams, R.M. and Sickles, R.C. (2007), Semi-parametric efficient distribution free estimation of panel models, *Communications in Statistics: Theory and Methods*, 36, 2425-2442.

Ahn, S.C., Good, D.H., and Sickles, R.C. (2000), Estimation of long-run inefficiency levels: a dynamic frontier approach, *Econometric Review*, 19, 461-492.

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

Ait-Sahalia, Y. and Duarte, J. (2003), Nonparametric option pricing under shape restrictions, *Journal of Econometrics*, 116, 9-47.

Almanidis, P, Qian, J, and Sickles, R.C. (2014), Stochastic frontiers with bounded inefficiency. Festschrift in honor of Peter Schmidt: Econometric Methods and Applications. Berlin: Springer Science and Business Media.

Amsler, C., Prokhorov, A., and Schmidt, P. (2016), Endogenous stochastic frontier models, *Journal of Econometrics*, 190, 280-288.

Amsler, C., Prokhorov, A., and Schmidt, P. (2017), Endogenous environmental variables in stochastic frontier models, *Journal of Econometrics*, 199, 131-140.

Assaf, A.G., Gillen, D., and Tsionas, E.G. (2014), Understanding relative efficiency among airports: a general dynamic model for distinguishing technical and allocative efficiency, *Transportation Research Part B*, 70, 18-34.

Atkinson, S.E. and Halvorsen, R. (1984), Parametric efficiency tests, economies of scale, and input demand in U.S. electric power generation, *International Economic Review*, 25, 647-662.

Barnett, W.A. (2002), Tastes and technology: Curvature is not sufficient for regularity, *Journal of Econometrics*, 108, 199-202.

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

Behrman, J.R., Lovell, C.K., Pollak, R.A., and Sickles, R.C. (1992), The CET-CES-generalized Leontief variable profit function: an application to Indian agriculture, *Oxford Economic Papers*, 44, 341-354.

Berger, A.N. and Mester, L.J. (1997), Inside the black box: What explains differences in the efficiencies of financial institutions, *Journal of Banking & Finance*, 21, 895-947.

Blundell, R., Horowitz, J.L., and Parey, M. (2012), Measuring the price responsiveness of gasoline demand: Economic shape restrictions and nonparametric demand estimation, *Quantitative Economics*, 3, 29-51.

Bos, J.W.B. and Koetter, M. (2009), Handling losses in translog profit models, *Applied Economics*, 43, 307-312.

Caudill, S.B. and Ford, J. (1993), Biases in frontier estimation due to heteroscedasticity, *Economics Letters*, 41, 17-20.

Christensen, L.R. and Jorgenson, D.W., and Lau, L.J. (1971), Conjugate duality and the transcendental logarithmic production function, *Econometrica*, 39, 255-56.

Caves, D., Christensen, L., and Tretheway, M. (1980), Flexible cost functions for multiproduct firms, *Review of Economics and Statistics,* 62, 477-481.

Chambers, R., Färe, R., Grosskopf, S., and Vardanyan, M. (2013), Generalized quadratic revenue functions, *Journal of Econometrics*, 173, 11-21.

Clark, J.A. (1996) Economic cost, scale efficiency, and competitive viability in banking, *Journal of Money, Credit and Banking*, 28, 342-364.

Cornwell, C., Schmidt, P., and Sickles, R.C. (1990), Production frontiers with cross-sectional and time-series variation in efficiency levels, *Journal of Econometrics*, 46, 185-200.

Desli, E., Ray, S.C., and Kumbhakar, S.C. (2003), A dynamic stochastic frontier production model with time-varying efficiency, *Applied Economics Letters*, 10, 623-626.

Diewert, W.E. (1971), An application of the Shephard Duality Theorem: a generalized Leontief production function, *Journal of Political Economy*, 79, 481-507.

Diewert, W.E. (1973), Functional forms for profit and transformation functions, *Journal of Economic Theory*, 6, 284-316.

Diewert, W.E. (1974a), A note on aggregation and elasticities of substitution, *Canadian Journal of Economics*, 7, 12-20.

Diewert, W.E. (1974b), Functional forms for revenue and factor requirements functions, *International Economic Review*, 15, 119-130.

Diewert, W.E. and Wales, T. J. (1987), Flexible functional forms and global curvature conditions, *Econometrica,* 55, 43-68.

Du, P., Parmeter, C.F., and Racine, J.S. (2013), Nonparametric kernel regression with multiple predictors and multiple shape constraints, *Statistica Sinica*, 23, 1347-71.

Duncombe, W. and Yinger, J. (2011), Making do: State constraints and local responses in California's education finance system, *International Tax and Public Finance*, 18, 337-368.

Duygun, M., Kutlu, L., and Sickles, R.C. (2016), Measuring productivity and efficiency: a Kalman filter approach, *Journal of Productivity Analysis*, 46, 155-167.

Fan, Y. and Li, Q., and Weersink A (1996), Semiparametric estimation of stochastic production frontier, *Journal of Business and Economic Statistics*, 14, 460-468.


Feng, G. and Serletis, A. (2008), Productivity trends in U.S. manufacturing: Evidence from the NQ and AIM cost functions, *Journal of Econometrics*, 142, 281-311.

Feng, G. and A. Serletis (2009), Efficiency and productivity of the U.S. banking industry, 1998-2005: Evidence from the Fourier cost function satisfying global regularity conditions, *Journal of Applied Econometrics*, 24, 105-138.

Fox, K.J. (1998), Non-parametric estimation of technical progress. *Journal of Productivity Analysis*, 10, 235-250.

Fuss, M. and McFadden, D. (1978), Production economics: a dual approach to theory and applications (II): applications of the theory of production, vol2, McMaster University Archive for the History of Economic Thought.

Gagnepain, P. and Ivaldi, M. (2002), Stochastic frontiers and asymmetric information models, *Journal of Productivity Analysis,* 18, 145-159.

Gagnepain, P. and Ivaldi, M. (2017), Economic efficiency and political capture in public service contracts, *Journal of Industrial Economics*, 65, 1-38.

Geman, S. and Geman, D. (1984), Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 609-628.

Getachew, L. and Sickles, R.C. (2007), The policy environment and relative price efficiency of Egyptian private sector manufacturing: 1987/88-1995/96, *Journal of Applied Econometrics*, 22, 703-728.

Gong, B. and Sickles, R.C. (2017), Resource allocation in multidivisional multiproduct firms: examining the divisional productivity of energy companies, *Working paper.*

Good, D.H., Nadiri, M.I., and Sickles, R.C. (1991), The structure of production, technical change and efficiency in a multiproduct industry: an application to U.S. airlines, *NBER Working paper* 3939.

Greene, W.H. (1980a), Maximum likelihood estimation of econometric frontier functions, *Journal of Econometrics*, 3, 27-56.

Greene, W.H. (1980b), On the estimation of a flexible frontier production model, *Journal of Econometrics*, 13, 101-115.

Greene, W.H. (2003), Simulated likelihood estimation of the normal-gamma stochastic frontier function, *Journal of Productivity Analysis*, 19, 179-190.

Greene, W.H. (2005a), Fixed and random effects in stochastic frontier models, *Journal of Productivity Analysis*, 23, 7-32.

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

Griffiths, W.E., O'Donnell, C. J., and Tan Cruz, A. (2000), Imposing regularity conditions on a system of cost and cost-share equations: A Bayesian approach, *Australian Journal of Agricultural and Resource Economics*, 44, 107-127.

Griffiths, W.E. and Hajargasht, G. (2016), Some models for stochastic frontiers with endogeneity, *Journal of Econometrics*, 190, 341-348.

Gronberg, T.J., Jansen, D.W., and Taylor, L.L. (2011), The adequacy of educational cost functions: Lessons from Texas, *Peabody Journal of Education*, 86, 3-27.

Guan Z., Kumbhakar S.C., Myers R.J., and Lansink A.O. (2009), Measuring excess capital capacity in agricultural production, *American Journal of Agricultural Economics*, 91, 765-776.

Guilkey, D.K. and Lovell, C.A.K. (1980), On the flexibility of the translog approximation, *International Economic Review*, 21. 137-147.

Guilkey, D.K., Lovell, C.A.K., and Sickles, R.C. (1983), A comparison of the performance of three flexible functional forms, *International Economic Review*, 24, 591-616.

Hall, P. and Huang, L.S. (2001), Nonparametric kernel regression subject to monotonicity constraints, *Annals of Statistics*, 29, 624-647.

Hastings, W.K. (1970), Monte Carlo sampling methods using Markov chains and their applications, *Biometrica*, 57, 97-109.

Henderson, D.J., List, J.A., Millimet, D.L., Parmeter, C.F., and Price, M.K. (2012), Empirical implementation of nonparametric first-price auction models. *Journal of Econometrics*, 168, 17-28.

Huang, T.H. and Chen, Y.H. (2009), A study on long-run inefficiency levels of a panel dynamic cost frontier under the framework of forward-looking rational expectations, *Journal of Banking and Finance*, 33, 842-849.

Hughes, J.P., Lang, W., Mester, L.J., and Moon, C.-G. (1995), Recovering banking technologies when managers are not risk-neutral, Federal Reserve Bank of Chicago Proceedings.

Hughes, J.P. and Mester, L.J. (1993), A quality and risk-adjusted cost function for banks: Evidence on the "too-big-to-fail" doctrine, *Journal of Productivity Analysis*, 4, 293-315.

Humphrey, D. and Pulley L. (1997), Banks' responses to deregulation: profits, technology, and efficiency, *Journal of Money, Credit and Banking*, 29, 73-93.

Jondrow, J., Lovell, C.A.K., Materov, I.S., and Schmidt, P. (1982), On the estimation of technical inefficiency in stochastic frontier production models, *Journal of Econometrics*, 19, 233-238.

Karakaplan, M.U., and Kutlu, L. (2017a), Handling endogeneity in stochastic frontier analysis, *Economics Bulletin, 37, 889-901*.

Karakaplan, M.U. and Kutlu, L. (2017b), Endogeneity in panel stochastic frontier models: An application to the Japanese cotton spinning industry, *Applied Economics*, 49, 5935-5939.

Karakaplan, M.U. and Kutlu, L. (2018), School district consolidation policies: Endogenous cost inefficiency and saving reversals, Forthcoming in *Empirical Economics*.

Kasman, A. and Yildirim, C. (2006), Cost and profit efficiencies in transition banking: The case of new EU members, *Applied Economics*, 38, 1079-1090.

Kleit, A., and Terrell, D. (2001), Measuring potential efficiency gains from deregulation of electricity generation: A Bayesian approach, *Review of Economics and Statistics*, 83, 523-530.

Koetter, M., Kolari, J.W., and Spierdijk, L. (2012), Enjoying the quiet life under deregulation? Evidence from adjusted Lerner indices for US banks, *Review of Economics and Statistics*, 94, 462-480.

Koop, G., Osiewalski, J., and Steel, M. (1997), Bayesian efficiency analysis through individual effects: Hospital cost Frontiers, *Journal of Econometrics*, 76, 77-105.

Kumbhakar, S.C. (1990), Production frontiers, panel data, and time-varying technical inefficiency, *Journal of Econometrics*, 46, 201-211.

Kumbhakar, S.C. (1994), A multiproduct symmetric generalized McFadden cost function, *Journal of Productivity Analysis,* 5, 349-357.

Kumbhakar, S.C. (1996), Efficiency measurement with multiple outputs and multiple inputs, *Journal of Productivity Analysis*, 7, 225-255.

Kumbhakar, S.C. (2001), Estimation of profit functions when profit is not maximum, *American Journal of Agricultural Economics*, 83, 1-19.

Kumbhakar, S.C. and Lai, H.P. (2016), Maximum likelihood estimation of the revenue function system with output-specific technical efficiency. *Economics Letters*, 138, 42-45.

Kuosmanen, T., and Kortelainen, M. (2012) Stochastic non-smooth envelopment of data: semi-parametric frontier estimation subject to shape constraints, *Journal of Productivity Analysis*, 38, 11-28.

Kutlu, L. (2010), Battese-Coelli estimator with endogenous regressors, *Economics Letters*, 109, 79-81.

Kutlu, L. and McCarthy, P. (2016), US airport governance and efficiency, *Transportation Research Part E*, 89, 117-132.

Kutlu, L. and Sickles, R.C. (2012), Estimation of market power in the presence of firm level inefficiencies, *Journal of Econometrics*, 168, 141-155.

Kutlu, L., Tran, K., and Tsionas, E.G. (2018), A time-varying true individual effects model with rndogenous tegressors, *Working paper*. Available at SSRN: https://ssrn.com/abstract=2721864.

Kutlu, L. and Wang, R. (2018), Estimation of cost efficiency without cost data, *Journal of Productivity Analysis*, 49, 137-151.

Lau, L. J. (1978), Testing and imposing monotonicity, convexity, and quasiconcavity. In: Fuss, M., McFadden, D. (eds.), *Production Economics, A Dual Approach to Theory and Applications*. Vol. (1978). Amsterdam: North-Holland, pp. 409-453.

Lee, Y.H. and Schmidt P. (1993), A production frontier model with flexible temporal variation in technical efficiency, in Fried H.O. and Schmidt S.S. (eds.) The Measuring Productivity Efficiency: Techniques and Applications, Oxford, U.K., 237-255.

Lewbel, A. (2010), Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models, *Journal of Business & Economic Statistics*, 30, 67-80.

Lovell, C.A.K. and Sickles, R.C. (1983), Testing efficiency hypothesis in joint production: a parametric approach, *Review of Economics and Statistics*, 65, 51-58.

Ma, S. and Racine, J.S. (2013), Additive regression splines with irrelevant categorical and continuous regressors, *Statistica Sinica*, 23(2), 515-541.

Magnus, J. (1979). Substitution between energy and non-energy inputs in the Netherlands, 1950-1976, *International Economic Review*, 20, 465-484.

Mairesse, J. and Jaumandreu J. (2005), Panel-data estimates of the production function and the revenue function: what difference does it make? *The Scandinavian Journal of Economics*, 107, 651-672.

Mammen, E. (1991), Estimating a smooth monotone regression function. *Annals of Statistics*, 19, 724-740.

Mammen, E. and Thomas-Agnan, C. (1999), Smoothing splines and shape restrictions, *Scandinavian Journal of Statistics: Theory and Applications*, 26, 239-252.

Matzkin, R.L. (1991), A nonparametric maximum rank correlation estimator. In: Barnett, W., Powell, J., Tauchen, G. (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. Cambridge: Cambridge University Press.

Matzkin, R.L. (1994), Restrictions of economic theory in nonparametric methods. In: Engel, R.F.& McFadden, D.L. (eds.), *Handbook of Econometrics*. Vol. IV. Amsterdam: Elsevier, Ch. 42.

Maudos, J., Pastor, J.M., Pérez, F., and Quesada, J. (2002), Cost and profit efficiency in European Banks, *Journal of International Financial Markets, Institutions and Money*, 12, 33-58.

Meeusen, W. and Van Den Broeck J. (1977), Efficiency estimation from Cobb-Douglas production function with composed errors, *International Economic Review*, 18, 435-444.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. (1953). Equations of state calculations by fast computing machines, *Journal of Chemical Physics,* 21, 1087-1092.

Mukerjee, H. (1988), Monotone nonparametric regression, *Annals of Statistics*, 16, 741-750.

Mutter, R.L., Greene, W.H., Spector, W., Rosko, M.D., and Mukamel, D.B. (2013), Investigating the impact of endogeneity on inefficiency estimates in the application of stochastic frontier analysis to nursing homes, *Journal of Productivity Analysis*, 39, 101-110.

Oliveira, R., Pedro, M.I., and Marques, R.C. (2013), Efficiency performance of the Algarve hotels using a revenue function, *International Journal of Hospitality Management*, 35, 59-67.

Perlof, J.M., Karp, L.S., and Golan, A. (2007), Estimating market power and strategies. Cambridge University Press.

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

Polak, R.A., Sickles, R.C., and Wales, T.J. (1984), The CES-Translog: Specification and estimation of a new cost function, *Review of Economics and Statistics*, 66, 602-607.

Rask, K. (1995), The structure of technology in Brazilian sugarcane production, 1975-87: An application of a modified symmetric generalized McFadden cost function, *Journal of Applied Econometrics*, 10, 221-232.

Ramsay, J.O. (1998), Estimating smooth monotone functions, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 60, 365-375.

Restrepo-Tobón, D. and Kumbhakar, S. (2014), Enjoying the quiet life under deregulation? Not quite, *Journal of Applied Econometrics*, 29, 333-343.

Rogers, K.E. (1998), Nontraditional activities and the efficiency of US commercial banks, *Journal of Banking & Finance*, 22, 467-482.

Röller, L.-H. (1990), Proper quadratic cost functions with an application to the Bell System, *Review of Economics and Statistics*, 72, 202-210.

Ruud, P.A. (1997), Restricted least squares subject to monotonicity and concavity constraints. In: Kreps, D.M.& Wallis, K.F. (eds.), *Advances in Economics and Econometrics: Theory and Applications, Proceedings of the Seventh World Congress.* Vol. 3. Cambridge: Cambridge University Press, pp. 166-187.

Ryan, D.L. and Wales, T.J. (2000), Imposing local concavity in the translog and generalized Leontief cost functions, *Economics Letters*, 67, 253-260.

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

Serletis, A. and Feng, G. (2015), Imposing theoretical regularity on flexible functional forms, *Econometric Reviews*, 34, 198-227.

Shively, T.S., Walker, S.G. & Damien, P. (2011), Nonparametric function estimation subject to monotonicity, convexity and other shape constraints, *Journal of Econometrics*, 161, 166-181.

Sickles, R.C., Good, D., and Johnson, R.L. (1986), Allocative distortions and the regulatory transition of the U.S. airline industry, *Journal of Econometrics*, 33, 143-163.

Sickles, R.C. and Streitwieser, M.L. (1998), An analysis of technology, productivity, and regulatory distortion in the interstate natural gas transmission industry: 1977-1985, *Journal of Applied Econometrics*, 13, 377-395.

Sickles, R. C. and Zelenyuk, V. (2018), Measurement of Productivity and Efficiency: Theory and Practice, with Valentin Zelenyuk. New York: Cambridge University Press, in process, 2018.

Simar, L., Van Keilegom, I. and Zelenyuk, V. (2017), Nonparametric least squares methods of stochastic frontier models, *Journal of Productivity Analysis*, 47, 189-204.

Stevenson, R. (1980), Likelihood functions for generalized stochastic frontier functions, *Journal of Econometrics*, 13, 57-66.

Terrell, D. (1996), Incorporating monotonicity and concavity conditions in flexible functional forms, *Journal of Applied Econometrics*, 11, 179-194.

Tran, K.C. and Tsionas, E.G. (2013), GMM estimation of stochastic frontier model with endogenous regressors, *Economics Letters,* 118, 233-236.

Tsionas, E.G. (2006), Inference in dynamic stochastic frontier models, *Journal of Applied Econometrics*, 21, 669-676.

Vander Vennet, R. (2002), Cost and profit efficiency of financial conglomerates and universal banks in Europe, *Journal of Money, Credit, and Banking*, 34, 254-282.

Varian, H. (1992), Microeconomic Analysis. W.W. Norton, New York.

Wales, T.J. (1977), On the flexibility of flexible functional forms: An empirical approach, *Journal of Econometrics*, 5, 183-193.

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

Weiher, J.C., Sickles, R.C. and Perloff, J.M. (2002), Market power in the US airline industry. Measuring Market Power (2002), Daniel J. Slottje, Ed., North-Holland, Emerald Group Publishing.

Wiley, D.E., Schmidt, W.H., and Bramble, W.J. (1973), Studies of a class of covariance structure models, *Journal of the American Statistical Association*, 68, 317-323.

Wu, X., and Sickles, R.C. (2018), Semiparametric estimation under shape constraints, *Econometrics and Statistics*, 6, 74-89.

Zellner, A., Kmenta, J., and Drèze, J. (1966), Specification and estimation of Cobb-Douglas production function models, *Econometrica*, 34, 784-795.