5 The Estimation and Decomposition of Profit Efficiency

5.1 INTRODUCTION

In Chapter 4 our efficiency analysis was based on the assumption that producers face exogenously determined input prices and output quantities and attempt to allocate inputs so as to minimize the cost of producing their outputs. Under this assumption inputs, but not outputs, are determined endogenously. While a cost minimization objective is undoubtedly appropriate in some environments, it can be argued that in other environments it is not sufficiently stringent, because for many producers the ultimate objective is to maximize profit. In such an environment the assumption is that producers face exogenously determined input prices and output prices and attempt to allocate inputs and outputs so as to maximize profit. Under this assumption both inputs and outputs are determined endogenously. The main difference between the two frameworks is that when producers attempt to maximize profit, they have to decide not only how much of various inputs to use (which is the case for producers attempting to minimize cost), but also how much of various outputs to produce. In other words, the issue is not just one of finding the cost-minimizing input combination required to produce a given bundle of outputs, but also one of finding the revenue-maximizing output combination as well, and so the number of decision variables increases from N to $M + N$. In this chapter our focus is on efficiency measurement when the objective of producers is to maximize profit.

A natural question to ask in a price-taking, profit-maximizing framework is: Can inefficient producers survive? The answer is: Not for long. In a long-run competitive equilibrium context profit is driven to zero and only efficient producers survive. However in a short-run temporary equilibrium context inefficient producers can survive with some loss, even in a competitive environment, provided that the loss is less than the cost of their fixed inputs. These considerations suggest that if one is interested in estimating profit efficiency in a price-taking environment, then it is appropriate to conduct the exercise within a short-run framework in which some inputs are exogenously determined. Thus the appropriate standard against which to evaluate profit efficiency is the variable profit frontier.

The chapter is organized as follows. In Section 5.2 we consider a situation in which profit-seeking producers produce a single output, and we introduce models based on primal production frontiers in Section 5.2.1 and dual variable profit frontiers in Section 5.2.2. Throughout this section we assume that output price, variable input prices, and quasi-fixed input quantities are exogenously determined. We show how to model technical and allocative inefficiencies, and we derive their impacts on variable profit. We discuss estimation procedures for primal and dual models in both cross-sectional and panel data settings.

In Section 5.3 we relax the single-output assumption and allow profit-seeking producers to produce multiple outputs. These multiple-output models are discussed within a primal distance function framework in Section 5.3.1 and a dual profit frontier framework in Section 5.3.2. We continue to assume that output prices and variable input prices are exogenously determined, and for purely econometric reasons to be explained later we ignore quasi-fixed inputs. We show how to model technical and allocative inefficiencies, and we derive their impacts on profit. We discuss estimation procedures in the context of both cross-sectional and panel data settings.

In Section 5.4 we consider a situation in which producers seek to maximize profit, but they no longer take output prices as exogenously determined. In contrast to the conventional scenarios considered in Sections 5.2 and 5.3, in which prices are assumed to be exogenously determined, perhaps by the forces of competition, in this section we endow producers with some pricing power in their product markets.

The result is an "alternative" profit frontier in which variable input prices and output quantities (rather than output prices) are exogenously determined. Because output prices are no longer taken to be exogenous, the alternative profit frontier incorporates product demand structure as well as technology structure, and so does not provide a dual representation of the structure of production technology. Nonetheless the central problem of modeling and estimating technical and allocative inefficiencies remains unchanged. However since the alternative profit frontier has different regularity conditions than a conventional profit frontier does, estimation procedures are modified. Here, as in Sections 5.2 and 5.3, our discussion of estimation procedures is brief, because only minor modifications to the procedures developed in Chapters 3 and 4 are required.

Section 5.5 contains references to the relevant literature.

5.2 SINGLE-OUTPUT MODELS

In this section we consider a situation in which producers produce a single output. In Section 5.2.1 we use a primal production frontier approach, in which the production frontier and the first-order conditions for variable profit maximization are used to estimate the parameters of the model as well as the magnitudes of technical and allocative inefficiencies and variable profit inefficiency. (Although this material is based on a production frontier, and hence is referred to as a primal approach, it also exploits dual price information, and so it is not really a pure primal approach.) This is followed in Section 5.2.2 with a dual variable profit frontier approach, in which one can either estimate a variable profit frontier using a single-equation method or estimate a system of equations consisting of either the variable profit frontier and the associated variable profit share equations or just the variable profit share equations. The central problem remains one of modeling and estimating technical and allocative inefficiencies and their impact on variable profit.

5.2.1 The Primal Production Frontier Approach

One implicit assumption in the production frontier approach adopted in Chapter 3 is that inputs are exogenous and uncorrelated with technical inefficiency. This assumption seems quite strong, since Mundlak (1961) and others have claimed that more efficient producers tend to produce more output and to use more inputs, given the input prices and the output prices they face. If so, technical inefficiency will be correlated with input use, resulting in inconsistent parameter estimates when a single-equation production frontier model is used. This problem can be alleviated by treating variable inputs as endogenous, which allows correlation between variable inputs and technical inefficiency.

We begin by writing the production frontier as

$$y = f(x, z; \beta)\cdot \exp\{-u\}, \tag{5.2.1}$$

where $y \geq 0$ is scalar output, $x = (x_1, \ldots, x_N) \geq 0$ is a vector of variable inputs, $z = (z_1, \ldots, z_Q) \geq 0$ is a vector of quasi-fixed inputs, $u \geq 0$ represents output-oriented technical inefficiency, and $f(x, z; \beta)$ is the deterministic kernel of a stochastic production frontier. If producers attempt to maximize variable profit (conditional on u), the first-order conditions can be written as

$$f_n(x, z; \beta)\cdot \exp\{-u\} = \frac{w_n}{p}\cdot \exp\{-\xi_n\}, \qquad n = 1, \ldots, N, \tag{5.2.2}$$

where $f_n(x, z; \beta) = \partial f(x, z; \beta)/\partial x_n$, the w_n/p are normalized variable input prices, $w = (w_1, \ldots, w_N) > 0$ is an input price vector, and $p > 0$ is the scalar output price. The ξ_n are interpreted as allocative inefficiencies, nonzero values of which indicate over- or underutilization of a variable input, given normalized prices of the variable inputs and quantities of the quasi-fixed inputs. Such allocative inefficiencies may also arise due to the presence of constraints (other than the production technology) faced by the producer that are not incorporated into the optimization problem. The effect of all of these influences will be captured by the ξ_n, which we call allocative inefficiencies.

The definition of allocative inefficiency in equation (5.2.2) is different from that used in Chapter 4. In Chapter 4 we defined input allocative inefficiencies as departures of marginal rates of substitution (MRS) from the respective input price ratios, whereas in equation (5.2.2) they are defined as departures of variable input marginal products from their normalized prices. The difference is due to the fact that in Chapter 4 we used a cost minimization criterion for which the first-order conditions are expressed by the equality of MRSs with input price ratios. Thus it was natural to define input allocative inefficiencies as departures from this rule. In a variable profit-maximizing framework the first-order conditions are different, but we follow the same principle and define variable input allocative inefficiencies in terms of nonfulfillment of the first-order conditions for variable profit maximization. These two measures are, however, related. The definition used in the cost-minimizing framework is provided by the ratios $\xi_{n1} = \xi_n/\xi_1$ for variable input pairs (x_n, x_1), $n = 2, \ldots, N$.

Cobb–Douglas If the production frontier takes the Cobb–Douglas form, the production frontier in equation (5.2.1) and the first-order conditions for variable profit maximization in equation (5.2.2) can be written in logarithmic form as

$$\ln y = \beta_o + \sum_n \beta_n \ln x_n + \sum_q \gamma_q \ln z_q + v - u, \tag{5.2.3}$$

$$\begin{aligned}\ln x_n &= \beta_o + \ln\beta_n + \sum_k \beta_k \ln x_k + \sum_q \gamma_q \ln z_q - \ln\frac{w_n}{p} - u + \xi_n,\\ n &= 1, \ldots, N,\end{aligned} \tag{5.2.4}$$

where v is the stochastic noise error component associated with the production frontier. In deriving the first-order conditions one has to start with either a deterministic production frontier or a stochastic production frontier evaluated at $v = 0$, or take the median value of output [with the assumption that v is distributed normally with mean 0 as in Kumbhakar (1987)]. The basic conclusion remains unchanged no matter how one treats the stochastic noise component v.

Solving the $(N + 1)$ equations (5.2.3) and (5.2.4) for the optimal values of the $(N + 1)$ endogenous variables gives the following output supply and variable input demand equations:

$$\begin{aligned}\ln y = {} & \frac{1}{1-r}\beta_o + \frac{1}{1-r}\sum_n \beta_n\left(\ln\beta_n - \ln\frac{w_n}{p}\right)\\ & + \frac{1}{1-r}\sum_q \gamma_q \ln z_q + \frac{1}{1-r}\sum_n \beta_n\xi_n - \frac{1}{1-r}u + v,\end{aligned} \tag{5.2.5}$$

$$\begin{aligned}\ln x_k = {} & \frac{1}{1-r}\beta_o + \frac{1}{1-r}\sum_n (\beta_n + (1-r)\delta_{nk})\left(\ln\beta_n - \ln\frac{w_n}{p}\right)\\ & + \frac{1}{1-r}\sum_q \gamma_q \ln z_q + \frac{1}{1-r}\sum_n (\beta_n + (1-r)\delta_{nk})\xi_n\\ & - \frac{1}{1-r}u, \qquad k = 1, \ldots, N,\end{aligned} \tag{5.2.6}$$

where $r = \Sigma_n\beta_n < 1$ and $\delta_{nk} = 1$ if $n = k$ and $\delta_{nk} = 0$ if $n \neq k$.

It can be seen from equations (5.2.5) and (5.2.6) that variable profit-maximizing output production and variable input use depend on normalized variable input prices and quasi-fixed input quantities. They also depend on the magnitudes of both technical and allocative inefficiencies. In particular, the relatively more technically efficient producers (those with $u \to 0^+$) use more of their variable inputs to produce more output, *ceteris paribus*. This conclusion supports the observation of Mundlak (1961), although it runs counter to the results obtained in Chapter 4. The reason is that input use in a cost minimization framework is conditional on exogenous outputs, whereas variable input use in the present variable profit maximization framework is unconditional. Moreover the orientation is different. In Chapter 4 technical efficiency was input oriented to maintain consistency with the cost minimization orientation, whereas here technical efficiency is output oriented, and relatively more technically efficient producers produce more output with more variable input use.

It is possible to estimate the system of output supply and variable input demand equations (5.2.5) and (5.2.6). However it is easier to estimate an alternative system consisting of the production frontier in equation (5.2.3) and the first-order conditions in equations (5.2.4), because the latter system has a more tractable error structure. Tractability comes at a cost, however, since equations (5.2.4) do not contain a random-noise error component, a structural characteristic we also encountered in Chapter 4. The distributional assumptions on the error components this model does contain are as follows:

(i) $v \sim \text{iid } N(0, \sigma_v^2)$.

(ii) $u \sim \text{iid } N^+(0, \sigma_u^2)$.

(iii) $\xi = (\xi_1, \ldots, \xi_N)' \sim \text{iid } N(0, \Sigma)$.

(iv) The elements of ξ are distributed independently of v and u, and v and u are distributed independently of one another.

With these distributional assumptions the log likelihood function for a sample of I producers can be written as

$$\ln L = \text{constant} - \frac{I}{2}\ln\sigma_v^2 - \frac{I}{2}\ln\sigma_u^2 - \frac{I}{2}\ln|\Sigma| + \frac{I}{2}\ln\sigma^2 - \frac{1}{2}\sum_i a_i + \sum_i \ln\Phi\left(-\frac{\mu_i}{\sigma}\right) + \frac{I}{2}\ln(1-r), \tag{5.2.7}$$

where $a_i = [\Xi_i^2/\sigma_v^2 + (b_i)'\Sigma^{-1}(b_i) - \sigma^2(\Xi_i/\sigma_v^2 + \iota'\Sigma^{-1}b_i)^2]$, $\sigma^2 = [1/\sigma_v^2 + 1/\sigma_u^2 + \iota'\Sigma^{-1}\iota]^{-1}$, $\mu_i = [\sigma^2(\Xi_i/\sigma_v^2 + \iota'\Sigma^{-1}b_i)]$, $\Xi_i = -u_i + v_i$, and $b_i = -\iota u_i + \xi_i$, ι being an $N \times 1$ vector of ones. The final term is the Jacobian of the transformation from $(\xi_i, b_{1i}, \ldots, b_{Ni})$ to $(\ln y_i, \ln x_{1i}, \ldots, \ln x_{Ni})$.

Maximization of this log likelihood function gives consistent and efficient estimates of all technology and inefficiency parameters. Once the parameters have been estimated, the technical inefficiency of each producer can be predicted from either the conditional mean or the conditional mode of u_i given (ξ_i, b_i'). These predictors are

$$E(u_i|\xi_i, b_i') = \mu_i + \sigma\frac{\phi(\mu_i/\sigma)}{\Phi(\mu_i/\sigma)}, \tag{5.2.8}$$

$$M(u_i|\xi_i, b_i') = \begin{cases} \mu_i & \text{if } \mu_i > 0, \\ 0 & \text{otherwise}, \end{cases} \tag{5.2.9}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are density and cumulative distribution functions of a standard normal variable. After producer-specific estimates of technical inefficiency have been obtained, producer-specific estimates of the allocative inefficiencies ξ_{ni} can be obtained by subtracting the estimates of either $E(u_i|\xi_i,b_i')$ or $M(u_i|\xi_i,b_i')$ from the residuals of equations (5.2.4). The generalization of this approach to panel data is straightforward, and is discussed in detail in Kumbhakar (1987).

Translog The methodology developed previously for estimating technical and allocative inefficiencies relative to a Cobb–Douglas production frontier can be extended to flexible functional forms. For example, consider a translog production frontier given by

$$\ln y = \beta_o + \sum_n \beta_n \ln x_n + \sum_q \gamma_q \ln z_q + \frac{1}{2}\sum_n\sum_k \beta_{nk}\ln x_n \ln x_k + \frac{1}{2}\sum_q\sum_r \gamma_{qr}\ln z_q \ln z_r + \sum_n\sum_q \delta_{nq}\ln x_n \ln z_q + v - u, \tag{5.2.10}$$

where $\beta_{nk} = \beta_{kn}\ \forall k \neq n$ and $\gamma_{qr} = \gamma_{rq}\ \forall r \neq q$. The first-order conditions for variable profit maximization, evaluated at $v = 0$, can be derived from equations (5.2.2) and (5.2.10) as

$$S_n = \varepsilon_n \cdot \exp\{\xi_n\} = \left[\beta_n + \sum_k \beta_{nk}\ln x_k + \sum_q \delta_{nq}\ln z_q\right] \cdot \exp\{\xi_n\}, \quad n = 1, \ldots, N, \tag{5.2.11}$$

where $S_n = w_n x_n/py$ is the ratio of expenditure on the nth variable input to total revenue, and $\varepsilon_n = \partial\ln f(x;z;\beta)/\partial\ln x_n$ is the elasticity of $f(x,z;\beta)$ with respect to the nth variable input. The random-noise error component in equation (5.2.10) is treated in the same manner as it was in the Cobb–Douglas model. Since y, the x_n, and the S_n are all endogenous, it is necessary to rewrite the share equations (to make the number of endogenous variables equal to the number of equations) as

$$\ln x_n = \ln y_{|v=0} + \ln\frac{w_n}{p} - \ln\left[\beta_n + \sum_k \beta_{nk}\ln x_k + \sum_q \delta_{nq}\ln z_q\right] + \xi_n, \quad n = 1, \ldots, N, \tag{5.2.12}$$

where the term $\ln y_{|v=0}$ on the left-hand side of equation (5.2.12) is replaced with the deterministic component of its translog expression on the right-hand side of equation (5.2.10). Equations (5.2.10) and (5.2.12) constitute a system of $N + 1$ equations to be estimated. This system collapses to the Cobb–Douglas system given in equations (5.2.3) and (5.2.4) if all second-order parameters are zero. Under the usual assumptions made on the error terms v, u, and ξ, the likelihood function can be derived and consistent estimates of all parameters can be obtained, exactly as in the structurally simpler Cobb–Douglas case. Once the parameters have been estimated the JLMS decomposition technique can be used to obtain producer-specific estimates of technical inefficiency. Producer-specific estimates of the allocative inefficiencies ξ_{ni} can be obtained from the residuals of equation (5.2.12), exactly as in the Cobb–Douglas case. Kumbhakar (1994) provides details.

We conclude this section with an observation concerning the primal approach to the estimation of variable efficiency. We argued in Section 5.1 that if one is interested in estimating profit efficiency in a price-taking environment, it is appropriate to do so within a short-run variable profit-seeking framework. We have done exactly that in this section. However by now it should be clear that specifying a short-run framework in the primal approach raises no new econometric issues. All that happens is that a vector of quasi-fixed inputs is added to the equation system to be estimated; the equations themselves remain unchanged. As with the case of a variable cost frontier in Section 4.2.1.3, the issues raised by quasi-fixed inputs are economic rather than econometric.

5.2.2 The Dual Variable Profit Frontier Approach

Since both types of inefficiency reduce variable profit, it is important to quantify the effect of these inefficiencies on variable profit. To this end we focus directly on the dual variable profit frontier. We begin by deriving some general results, and then we consider Cobb–Douglas and translog functional forms.

We begin by allowing for technical inefficiency only. Corresponding to the production frontier specified in equation (5.2.1) is the dual variable profit frontier

$$v\pi = v\pi(pe^{-u}, w, z;\beta) = v\pi(p, w, z;\beta)\cdot h(p, w, z, u, \beta), \tag{5.2.13}$$

where $v\pi = py - w^Tx = (pe^{-u})(ye^{u}) - w^Tx$. Maximum variable profit in the presence of output-oriented technical inefficiency is

$$v\pi(pe^{-u}, w, z;\beta) = \max_{ye^{u},x}\{pe^{-u}ye^{u} - w^Tx : ye^{u} = f(x,z;\beta)\},$$

and maximum variable profit in the presence of technical efficiency $(u = 0)$ is

$$v\pi(p,w,z;\beta) = \max_{y,x}\{py - w^Tx : y = f(x,z;\beta)\}.$$

Consequently $h(p,w,z,u,\beta) = v\pi(pe^{-u},w,z;\beta)/v\pi(p,w,z;\beta)$ is the ratio of maximum variable profit (allowing for technical inefficiency

but assuming allocative efficiency) to maximum variable profit. Thus there is an exact relationship between output-oriented technical inefficiency and variable profit inefficiency, with the functional form of $h(p,w,z,u,\beta)$ depending on the functional form of $v\pi(p,w,z;\beta)$. It is clear that $h(p,w,z,u,\beta) \le 1$ since $pe^{-u} \le p \Rightarrow v\pi(pe^{-u},w,z;\beta) \le v\pi(p,w,z;\beta)$. It is also clear that the impact of technical inefficiency on variable profit depends on more than just the magnitude of technical inefficiency, since $h(p,w,z,u,\beta)$ depends on (p,w,z) as well as on u. Consequently if one estimates a relationship of the form $\ln v\pi(pe^{-u},w,z;\beta) = \ln v\pi(p,w,z;\beta) + \alpha u$, $\alpha \le 0$, this relationship is misspecified unless the underlying technology satisfies certain restrictions sufficient to cause $\ln h(p,w,z,u,\beta) = \alpha u$.

A result of Lau (1978; 151) can be applied to the present context to show that $f(\lambda x, z;\beta) = \lambda^r\cdot f(x,z;\beta)$, $\lambda > 0 \Leftrightarrow \ln v\pi(pe^{-u},w,z;\beta) = \ln v\pi(p,w,z;\beta) + \ln h(p,w,z,u,\beta) = \{[(1/(1-r))\ln p + \ln g(w) + \ln k(z)] - (1/(1-r))u\}$, where $g(w)$ is homogeneous of degree $[-r/(1-r)]$. Consequently estimation of the preceding variable profit relationship generates biased and inconsistent parameter estimates unless the underlying technology is homogeneous of degree $0 < r < 1$ in x. The bias and inconsistency arise because $E[h(p,w,z,u,\beta)]$ depends on (p, w, z), which also appear in $v\pi(p,w,z;\beta)$. If technology is homogeneous in x, then $\ln h(p,w,z,u,\beta) = -[1/(1-r)]u$ is independent of (p, w, z), and so bias and inconsistency do not arise.

The previous results generalize easily to the case in which both technical and allocative inefficiencies are present. If technical inefficiency is output oriented as in equation (5.2.1), and if allocative inefficiencies are modeled as in equation (5.2.2), then the dual variable profit frontier can be expressed as

$$v\pi = v\pi(pe^{-u}, w^s, z;\beta) = v\pi(p,w,z;\beta)\cdot h(p,w,z,u,\beta,\xi), \tag{5.2.14}$$

where $w^s = (w_1^s,\ldots,w_N^s) = (w_1\cdot\exp\{-\xi_1\},\ldots,w_N\cdot\exp\{-\xi_N\})$, $v\pi(pe^{-u}, w^s, z;\beta)$ is maximum variable profit in the presence of both types of inefficiency, $v\pi(p,w,z;\beta)$ is maximum variable profit in the absence of both types of inefficiency, and $h(p,w,z,u,\beta,\xi) = v\pi(pe^{-u}, w^s z;\beta)/v\pi(p,w,z;\beta)$. Since $v\pi(pe^{-u},w^s,z;\beta) \le v\pi(p,w,z;\beta)$, $h(p,w,z,u,\beta,\xi) \le 1$. Thus the variable profit loss due to inefficiency is given by the function $h(p,w,z,u,\beta,\xi)$, whose functional form depends on that of $v\pi(p,w,z;\beta)$.

Unfortunately the function $h(p,w,z,u,\beta,\xi)$ is not generally separable, so that $h(p,w,z,u,\beta,\xi) \ne h_1(p,w,z,u,\beta)\cdot h_2(p,w,z,\beta,\xi)$, and it is not generally possible to decompose profit inefficiency $h(p, w, z, u, \beta, \xi)$ into the product of its technical inefficiency and allocative inefficiency components. However if the production technology is homogeneous of degree r in x, then $\ln v\pi(pe^{-u}, w^s, z; \beta) = \ln v\pi(p, w, z; \beta) + \ln h(p, w, z, u, \beta, \xi) = \ln v\pi(p, w, z; \beta) + \{-(1/(1-r))u + [\ln g(w^s) - \ln g(w)]\}$. Thus under homogeneity $h(p, w, z, u, \beta, \xi) = h_1(p, w, z, u, \beta) \cdot h_2(p, w, z, \beta, \xi)$ with $h_1(p, w, z, u, \beta) = h_1(u)$ and $h_2(p, w, z, \beta, \xi) = h_2(w, \xi)$. Once again ignoring technical inefficiency generates biased and inconsistent parameter estimates unless the underlying technology is homogeneous. Worse still, ignoring allocative inefficiency generates biased and inconsistent parameter estimates even if technology is homogeneous, since $h_2(w, \xi)$ is independent of p and z, but not of w. This adverse impact on parameter estimates also carries over to estimates of supply and demand elasticities and other characteristics of the structure of production technology. This calls into question the assertion of Schmidt (1985–1986; 320) that "[t]he only compelling reason to estimate production frontiers is to measure efficiency." An equally compelling reason is to avoid biased and inconsistent estimates of the structure of production technology, and consequently of efficiency measured relative to that technology.

Cobb–Douglas We start with the homogeneous Cobb–Douglas form for simplicity. To simplify the algebra, we begin with the stochastic production frontier and derive the first-order conditions, conditional on both v and u. Under these conditions $\beta_n(y/x_n) = (w_n/p)\exp\{-\xi_n\} \Rightarrow \ln x_n - \ln y = \ln \beta_n - \ln(w_n / p) + \xi_n$, $n = 1, \ldots, N$, and a little algebra yields the dual normalized variable profit frontier

$$\ln\frac{v\pi}{p} = \ln\left[\frac{v\pi(p,w,z;\beta)}{p}\right] + \ln v\pi_u + \ln v\pi_\xi + \ln v\pi_v, \tag{5.2.15}$$

where

$$\begin{aligned}\ln\frac{v\pi(p,w,z;\beta)}{p} &= \frac{1}{1-r}\beta_o + \frac{1}{1-r}\sum_n \beta_n \ln(w_n/p) \\ &\quad + \frac{1}{1-r}\sum_q \gamma_q \ln z_q + \ln(1-r)\end{aligned} \tag{5.2.16}$$

is the normalized variable profit frontier in the presence of technical and allocative efficiency, $\ln v\pi_u = -[1/(1-r)]u \le 0$ represents the impact of technical inefficiency on normalized variable profit, $\ln v\pi_\xi = (E - \ln r) \le 0$ represents the impact of allocative inefficiency on normalized variable profit, where $E = [(1/(1-r))\Sigma_n\beta_n\xi_n + \ln[1 - \Sigma_n\beta_n\exp\{\xi_n\}]]$, $\ln v\pi_v = [1/(1-r)]v$ represents the impact of statistical noise on normalized variable profit, and $r = \Sigma_n\beta_n < 1$ measures the degree of homogeneity of $f(x, z; \beta)$ in x.

The normalized variable profit frontier in equation (5.2.15) shows that the overall variable profit inefficiency (the percentage loss of normalized variable profit due to both technical and allocative inefficiencies, $[\ln v\pi_u + \ln v\pi_\xi]$) is independent of output prices, input prices, and quasi-fixed input quantities. Recall that under homogeneity alone, $\ln h_1(p, w, z, u, \beta) = \ln v\pi_u$ is independent of (p, w, z), and $\ln h_2(p, w, z, \beta, \xi) = \ln v\pi_\xi$ is independent of p and z, but not of w. However in the homogeneous Cobb–Douglas case $\ln h_2(p, w, z, \beta, \xi) = \ln v\pi_\xi$ is independent of p, w, and z.

If the normalized variable profit frontier in equation (5.2.15) is rewritten as

$$\ln\frac{v\pi}{p} = \delta_o + \sum_n \delta_n \ln\frac{w_n}{p} + \sum_q \delta_q \ln z_q + v_\pi + u_\pi, \tag{5.2.17}$$

where δ_o is a constant, $\delta_n = -(1/1-r)\beta_n\ \forall n$, the $\delta_q = [1/(1-r)]\gamma_q\ \forall q$, $v_\pi = [1/(1-r)]v$, and $u_\pi = [\ln \pi_u + \ln \pi_\xi] \le 0$ (which is the overall normalized variable profit inefficiency), then equation (5.2.17) is structurally similar to the stochastic production frontier model in equation (5.2.3). Thus no new tools are required to estimate it. The MLE techniques developed in Chapter 3 are appropriate, and after estimation the JLMS decomposition can be used to generate producer-specific estimates of overall normalized variable profit inefficiency. However if one wants to disentangle the separate effects of technical and allocative inefficiencies, a system approach is required. Kumbhakar (1987) provides the details and we will consider such a system in Chapter 6.

Translog I We begin by allowing for technical inefficiency, but we maintain the assumption of allocative efficiency. If the dual variable profit frontier takes the translog form, then

$$\begin{aligned}
\ln v\pi = \ln v\pi(pe^{-u}, w, z; \beta) &= \beta_o + \beta_p \ln(pe^{-u}) \\
&+ \sum_n \beta_n \ln w_n + \sum_q \gamma_q \ln z_q + \frac{1}{2}\beta_{pp}[\ln(pe^{-u})]^2 \\
&+ \frac{1}{2}\sum_n \sum_k \beta_{nk} \ln w_n \ln w_k + \frac{1}{2}\sum_q \sum_r \gamma_{qr} \ln z_q \ln z_r \\
&+ \sum_n \beta_{pn} \ln(pe^{-u}) \ln w_n + \sum_q \beta_{pq} \ln(pe^{-u}) \ln z_q \\
&+ \sum_n \sum_q \gamma_{nq} \ln w_n \ln z_q, \qquad (5.2.18)
\end{aligned}$$

where linear homogeneity in (p, w) is imposed by the restrictions $\Sigma_n \beta_n + \beta_p = 1$, $\Sigma_n \beta_{np} + \beta_{pp} = 0$, $\Sigma_k \beta_{nk} + \beta_{np} = 0\ \forall n$, $\Sigma_n \gamma_{nq} + \beta_{pq} = 0\ \forall q$. These restrictions can be embedded by writing equation (5.2.18) as a normalized variable profit frontier

$$\begin{aligned}
\ln \frac{v\pi}{p} = \beta_o &+ \sum_n \beta_n \ln \frac{w_n}{pe^{-u}} + \frac{1}{2}\sum_n \sum_k \beta_{nk} \ln \frac{w_n}{pe^{-u}} \ln \frac{w_k}{pe^{-u}} \\
&+ \sum_q \gamma_q \ln z_q + \frac{1}{2}\sum_q \sum_r \gamma_{qr} \ln z_q \ln z_r \\
&+ \sum_n \sum_q \gamma_{nq} \ln \frac{w_n}{pe^{-u}} \ln z_q - u, \qquad (5.2.19)
\end{aligned}$$

which is homogeneous of degree 0 in (pe^{-u}, w). This normalized variable profit frontier can be expressed as

$$\ln \frac{v\pi}{p} = \ln v\pi\left(\frac{w}{p}, z; \beta\right) + \ln h(p, w, z, u, \beta), \qquad (5.2.20)$$

where $\ln v\pi(w/p, z; \beta)$ is the translog normalized variable profit frontier in the presence of technical efficiency and

$$\begin{aligned}
\ln h(p,w,z,u,\beta) &= -\left[\beta_p + \beta_{pp} \ln p + \sum_n \beta_{pn} \ln w_n + \sum_q \gamma_{pq} \ln z_q\right] u \\
&+ \frac{1}{2}\beta_{pp} u^2 = -u\left[1 - \sum_n \beta_n - \sum_n \sum_k \beta_{nk} \ln \frac{w_n}{p}\right. \\
&\left. - \sum_n \sum_q \gamma_{nq} \ln z_q\right] + \frac{1}{2}u^2 \sum_n \sum_k \beta_{nk}, \qquad (5.2.21)
\end{aligned}$$

which does not collapse to $\ln h(u)$ unless the underlying technology is homogeneous ($\beta_{pp} = \beta_{pn} = \gamma_{pq} = 0\ \forall n, q$). Nonetheless, even in the nonhomogeneous translog case, we have an exact relationship between technical inefficiency u and normalized variable profit inefficiency $h(p, w, z, u, \beta)$, and this relationship holds even though the translog form is not self-dual.

Hotelling's lemma may be applied to the variable profit frontier in equation (5.2.18) to generate the actual variable profit share equations

$$-\frac{w_n x_n}{v\pi} = -S_n(pe^{-u}, w; \beta) = \beta_n + \sum_k \beta_{nk} \ln \frac{w_k}{pe^{-u}} + \sum_q \gamma_{nq} \ln z_q, \quad n = 1, \ldots, N, \qquad (5.2.22)$$

$$\frac{py}{v\pi} = S_p(pe^{-u}, w; \beta) = 1 + \sum_n S_n(pe^{-u}, w; \beta). \qquad (5.2.23)$$

Equation (5.2.21) shows that it is impossible to disentangle u from (p, w, z) in the variable profit frontier without sacrificing its flexibility. Consequently estimation is based on the input variable profit share equations, to which we append classical error terms and rewrite as

$$\frac{w_n x_n}{v\pi} = -\left[\beta_n + \sum_k \beta_{nk} \ln \frac{w_k}{p} + \sum_q \gamma_{nq} \ln z_q\right] - \beta_{np} u + \eta_n, \quad n = 1, \ldots, N, \qquad (5.2.24)$$

where $\Sigma_k \beta_{nk} = -\beta_{np}$. The system of share equations (5.2.24) does not include the intercept term β_o or the parameters γ_q and γ_{qr} associated with the quasi-fixed inputs. Estimation of these parameters is considered beneath equations (5.2.28) and (5.2.29).

In a cross-sectional framework estimation of producer-specific technical inefficiencies requires that u be random. Here we follow standard practice in the stochastic frontier literature by assuming that $u \sim \text{iid } N^+(0, \sigma_u^2)$, that $\eta = (\eta_1, \ldots, \eta_N)' \sim \text{iid } N(0, \Sigma)$, and that the elements of η are distributed independently of u. The justification for these assumptions is that u is under the control of producers but η is not.

Given a sample of I producers, we need to find the density function of the error vector $[-\beta_{1p}u_i + \eta_{1i}, \ldots, -\beta_{Np}u_i + \eta_{Ni}]'$, $i = 1, \ldots, I$, in equations (5.2.24). Define $\Xi_i = (bu_i + \eta_i)'$, where $b = -(\beta_{1o}, \ldots, \beta_{No})'$. Since both u_i and the η_{ni} are iid across producers, we temporarily delete the producer subscript. The density function of Ξ can be expressed as

$$f(\Xi) = \int_0^\infty f(\Xi, u)du = \int_0^\infty f(\Xi|u)h(u)\,du, \tag{5.2.25}$$

where $f(\Xi, u)$ is the joint density function of Ξ and u, and $h(u)$ is the density function of u. Using the distributional assumptions on u and η, the density function of Ξ can be expressed as

$$\begin{aligned} f(\Xi) &= \frac{2}{(2\pi)^{(N+1)/2}|\Sigma|^{1/2}\sigma_u}\int_0^\infty \exp\left\{-\frac{1}{2}\left[(\Xi - bu)'\Sigma^{-1}(\Xi - bu) + \frac{u^2}{\sigma_u^2}\right]\right\}du \\ &= \frac{2\sigma\exp\{-a/2\}}{(2\pi)^{N/2}|\Sigma|^{1/2}\sigma_u}\Phi(\Xi'\Sigma^{-1}b\sigma), \end{aligned} \tag{5.2.26}$$

where $\sigma^2 = (1/\sigma_u^2 + b'\Sigma^{-1}b)^{-1}$, $a = \Xi'\Sigma^{-1}\Xi - \sigma^2(\Xi'\Sigma^{-1}b)^2$, and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal variable. Reintroducing producer subscripts, the log likelihood function for a sample of I producers can be expressed as

$$\begin{aligned} \ln L &= \text{constant} - \frac{I}{2}\ln|\Sigma| + I\ln\sigma + \sum_i \Phi\left(\Xi_i'\Sigma^{-1}b\sigma\right) \\ &\quad - I\ln\sigma_u - \frac{1}{2}\sum_i a_i, \end{aligned} \tag{5.2.27}$$

maximization of which generates consistent estimates of all parameters of the variable profit function that are included in the share equations, as well as of σ_u^2 and the elements of Σ. After obtaining parameter estimates, producer-specific estimates of technical inefficiency can be obtained from the decomposition formula in Kumbhakar (1987), which generalizes the JLMS formula to a simultaneous-equation system. Thus

$$E(u_i|bu_i + \eta_i) = \mu_i + \sigma\frac{\phi(\mu_i/\sigma)}{\Phi(\mu_i/\sigma)}, \tag{5.2.28}$$

$$M(u_i|bu_i + \eta_i) = \begin{cases} \mu_i & \text{if } \mu_i > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{5.2.29}$$

where $\mu_i = \Xi_i'\Sigma^{-1}b\sigma^2$ and $\phi(\cdot)$ is the density function of a standard normal variable.

Once u has been estimated from either the conditional mean or the conditional mode, profit efficiency can be estimated from the expression for $\ln h(p, w, z, u, \beta)$ in the second line of equation (5.2.21), all parameters of which are contained in the input variable profit share equations (5.2.24). Thus in the present formulation it is possible to estimate both technical inefficiency u and the variable profit loss arising from technical inefficiency $h(p, w, z, u, \beta)$.

Estimates of the remaining parameters (β_o, the γ_q, and the γ_{qr}) can be obtained from a second-stage regression conditional on the parameter estimates obtained from the share equations and estimates of u from either equation (5.2.28) or (5.2.29). This second-stage regression uses the following residuals obtained from equation (5.2.19):

$$\begin{aligned} \ln\frac{v\pi}{p} - \Bigg[\sum_n \beta_n \ln\frac{w_n}{pe^{-u}} - \frac{1}{2}\sum_n\sum_k \beta_{nk}\ln\frac{w_n}{pe^{-u}}\ln\frac{w_k}{pe^{-u}} \\ - \sum_n\sum_q \gamma_{nq}\ln\frac{w^n}{pe^{-u}}\ln z_q - u\Bigg], \end{aligned}$$

which are regressed using OLS on an intercept, the $(\ln z_q)$, and the $(\ln z_q)$ $(\ln z_r)$. This second-stage regression will give consistent estimates of the remaining parameters in the normalized variable profit frontier given in equation (5.2.19).

If panel data are available, and if one is willing to assume that technical inefficiency is time invariant, then it is possible to estimate all parameters in the model, and u and $h(p, w, z, u, \beta)$ as well, without having to impose strong distributional assumptions. The estimation procedure is simple and straightforward. Imposing the homogeneity restrictions, the normalized variable profit frontier and the associated input variable profit share equations can be written as

$$\ln\frac{\pi_{it}}{p_{it}} = \beta_o + \sum_n \beta_n \ln\frac{w_{nit}}{p_{it}e^{-ui}} + \frac{1}{2}\sum_n\sum_k \beta_{nk}\ln\frac{w_{nit}}{p_{it}e^{-ui}}\ln\frac{w_{kit}}{p_{it}e^{-ui}}$$
$$+\sum_q \gamma_q \ln z_{qit} + \frac{1}{2}\sum_q\sum_r \gamma_{qr}\ln z_{qit}\ln z_{rit}$$
$$+\sum_n\sum_q \gamma_{nq}\ln\frac{w_{nit}}{p_{it}e^{-ui}}\ln z_{qit} - u_i + v_{it}, \tag{5.2.30}$$

$$S_{nit} = -\left[\beta_n + \sum_k \beta_{nk}\ln\frac{w_{kit}}{p_{it}e^{-ui}} + \sum_q \gamma_{nq}\ln z_{qit}\right] + \eta_{nit},$$
$$n = 1,\dots,N, \tag{5.2.31}$$

where v_{it} and the η_{nit} are stochastic noise components, and the u_i are assumed to be producer-specific fixed effects. Under the assumption that v and the η_n have zero means and constant variance–covariance matrix, the system of equations (5.2.30) and (5.2.31) can be estimated using a nonlinear ITSUR procedure. Technical inefficiencies can be estimated relative to the most efficient producer in the sample, as in Schmidt and Sickles (1984). Once the parameters and the u_i have been estimated, profit loss due to technical inefficiency can be estimated for each producer using equation (5.2.21).

It should be noted that technical inefficiency appears in the normalized variable profit frontier both additively as an error component and interactively with the normalized input prices, and so it also appears interactively in the input variable profit share equations. Consequently a specification that includes only producer-specific intercepts in the normalized variable profit frontier to capture the impact of technical inefficiency is misspecified and generates inconsistent parameter estimates.

Translog II We now allow for both technical and allocative inefficiency. Since actual variable profit $v\pi = py - w^Tx$, and since $v\pi(pe^{-u}, w^s, z; \beta) = pe^{-u}ye^{u} - w^{sT}x$, an alternative expression for actual variable profit can be obtained by combining these two expressions to yield

$$v\pi = v\pi(pe^{-u}, w^s, z;\beta) + \sum_n (w_n^s - w_n)x_n$$
$$= v\pi(pe^{-u}, w^s, z;\beta)\cdot\left[1 - \sum_n(\exp\{\xi_n\}-1)S_n^s\right], \tag{5.2.32}$$

where shadow input prices $w_n^s = w_n\cdot\exp\{-\xi_n\}\ \forall n$ and shadow input variable profit shares $S_n^s = -\partial\ln v\pi(pe^{-u}, w^s, z;\beta)/\partial\ln w_n^s\ \forall n$. Observed input variable profit shares can be expressed in terms of shadow input variable profit shares as

$$S_n = \frac{w_n x_n(pe^{-u}, w^s, z;\beta)}{v\pi}$$
$$= \left[\frac{w_n^s x_n(pe^{-u}, w^s, z;\beta)}{v\pi(pe^{-u}, w^s, z;\beta)}\right]\cdot\left[\frac{v\pi(pe^{-u}, w^s, z;\beta)}{v\pi}\right]\cdot\left[\frac{w_n}{w_n^s}\right]$$
$$= \frac{\exp\{\xi_n\}S_n^s}{\left[1-\Sigma_k(\exp\{\xi_k\}-1)S_k^s\right]}. \tag{5.2.33}$$

If $v\pi(pe^{-u}, w^s, z; \beta)$ takes the translog functional form, then

$$\ln v\pi(pe^{-u}, w^s, z;\beta)$$
$$= \beta_o + \beta_p\ln(pe^{-u}) + \sum_n \beta_n \ln w_n^s + \sum_q \gamma_q \ln z_q$$
$$+\frac{1}{2}\beta_{pp}[\ln(pe^{-u})]^2 + \frac{1}{2}\sum_n\sum_k \beta_{nk}\ln w_n^s \ln w_k^s$$
$$+\frac{1}{2}\sum_q\sum_r \gamma_{qr}\ln z_q \ln z_r + \sum_n \beta_{pn}\ln(pe^{-u})\ln w_n^s$$
$$+\sum_q \beta_{pq}\ln(pe^{-u})\ln z_q + \sum_n\sum_q \gamma_{nq}\ln w_n^s \ln z_q, \tag{5.2.34}$$

and the shadow input variable profit share equations become

$$S_n^s = -\frac{\partial\ln v\pi(pe^{-u}, w^s, z;\beta)}{\partial\ln w_n^s}$$
$$= -\left(\beta_n + \sum_k \beta_{nk}\ln\frac{w_k^s}{p} + \sum_q \gamma_{nq}\ln z_q - \sum_k \beta_{nk}u\right)$$
$$= A_n - \beta_{np}u, \qquad n = 1,\dots,N, \tag{5.2.35}$$

where the linear homogeneity restrictions listed beneath equation (5.2.18) have been imposed. Alternatively, if a normalized shadow variable profit function is specified, the homogeneity restrictions are automatically imposed. The normalized version of equation (5.2.34) is

$$\ln v\pi\left(\frac{w^s}{pe^{-u}}, z; \beta\right) = \beta_o + \sum_n \beta_n \ln \frac{w_n^s}{pe^{-u}}$$
$$+\frac{1}{2}\sum_n \sum_k \beta_{nk} \ln \frac{w_n^s}{pe^{-u}} \ln \frac{w_k^s}{pe^{-u}} + \sum_q \gamma_q \ln z_q$$
$$+\frac{1}{2}\sum_q \sum_r \gamma_{qr} \ln z_q \ln z_r + \sum_n \sum_q \gamma_{nq} \ln \frac{w_n^s}{pe^{-u}} \ln z_q, \tag{5.2.36}$$

and the expression for actual variable profit given in equation (5.2.32) becomes, in normalized form,

$$\ln \frac{v\pi}{p} = \ln v\pi\left(\frac{w^s}{pe^{-u}}, z; \beta\right) + \ln\left[1 - \sum_n (\exp\{\xi_n\} - 1) S_n^s\right] - u. \tag{5.2.37}$$

It is now possible to derive an expression for variable profit inefficiency in the presence of both technical and allocative inefficiency. The generalization of the expression for $\ln h(p, w, z, u, \beta)$ in equation (5.2.21) becomes

$$\begin{aligned}
\ln h(p,w,z,u,\beta,\xi) = &\Bigg[\Bigg[1 - \sum_n \beta_n - \sum_n \sum_k \beta_{nk} \ln \frac{w_k}{p} + \sum_q \gamma_{pq} \ln z_q \\
&+\frac{1}{2}\sum_n \sum_k \beta_{nk}(-u)\Bigg](-u)\Bigg] + \Bigg[\sum_n \beta_n(-\xi_n) \\
&+\frac{1}{2}\sum_n \sum_k \beta_{nk}(-\xi_n)(-\xi_k) \\
&+\sum_n \sum_k \beta_{nk}(-\xi_n)\ln \frac{w_k}{p} \\
&+\ln\left\{1 - \sum_n (\exp\{\xi_n\} - 1)(A_n)\right\}\Bigg] \\
&+\Bigg[\ln\left\{1 - \sum_n (\exp\{\xi_n\} - 1)(-\beta_{np} u)\right\} \\
&-u\sum_n \sum_k \beta_{nk}(-\xi_n)\Bigg] \\
= &\ln h_1(p,w,z,u,\beta) + \ln h_2(p,w,z,\beta,\xi) \\
&+ \ln h_3(p,w,z,u,\beta,\xi).
\end{aligned} \tag{5.2.38}$$

It is clear that overall variable profit inefficiency $\ln h(p, w, z, u, \beta, \xi)$ cannot be decomposed into a technical inefficiency term $\ln h_1(p, w, z, u, \beta)$ and an allocative inefficiency term $\ln h_2(p, w, z, \beta, \xi)$ unless the interaction term $\ln h_3(p, w, z, u, \beta, \xi) = 0$. This term is zero if the underlying production technology is homogeneous, and even then the separate effects of technical and allocative inefficiencies both depend on (p, w, z).

We now consider how to estimate the translog model in the presence of both technical and allocative inefficiency. We begin by deriving the actual input variable profit share equations, which in the translog case can be expressed in terms of the shadow input variable profit share equations as

$$\frac{S_n \exp\{-\xi_n\}[1 - \Sigma_k(\exp\{\xi_k\} - 1)A_k] - A_n}{\beta_{no} + \exp\{-\xi_n\} S_n \Sigma_k (\exp\{\xi_k\} - 1)\beta_{kp}} = -u, \qquad n = 1, \ldots, N. \tag{5.2.39}$$

Adding classical error terms η_n to equations (5.2.39) generates a system of N equations with error vector $-(\iota u + \eta)$, ι being an $N \times 1$ column vector of ones. Writing these equations in implicit form, we have

$$q_n\left(S_n, \frac{w}{p}, \xi\right) = -u + \eta_n, \qquad n = 1, \ldots, N, \tag{5.2.40}$$

or, in vector notation,

$$q\left(S, \frac{w}{p}, \xi\right) = -\iota u + \eta = \Xi. \tag{5.2.41}$$

The density function of Ξ is

$$f(\Xi) = \frac{2\sigma \exp\{-a/2\}}{(2\pi)^{N/2}|\Sigma|^{1/2}\sigma_u}\Phi(-\Xi'\Sigma^{-1}\iota\sigma), \tag{5.2.42}$$

where $\sigma^2 = (1/\sigma_u^2 + \iota'\Sigma^{-1}\iota)^{-1}$ and $a = \Xi'\Sigma^{-1}\Xi - \sigma^2(\Xi'\Sigma^{-1}\iota)^2$. The log likelihood function is

$$\ln L = \text{constant} - \frac{I}{2}\ln|\Sigma| + I \ln \sigma + \sum_i \ln \Phi(-\Xi_i'\Sigma^{-1}\iota\sigma)$$
$$- I \ln \sigma_u - \frac{1}{2}\sum_i a_i + \sum_i \ln|D_i|, \tag{5.2.43}$$

where D_i is the Jacobian of the transformation from $(\Xi_{1i}, \ldots, \Xi_{Ni})$ to $(S_{1i}, \ldots, S_{Ni})$. This log likelihood function can be maximized to obtain estimates of all technology parameters appearing in the share equations, as well as of σ_u^2 and the elements of the Σ matrix. Once the parameters have been estimated, producer-specific estimates of technical efficiency can be obtained from modifications of equations (5.2.28) and (5.2.29). The modifications consist of replacing b with $-\iota$ in the expressions for μ_i and σ^2.

Some of the remaining parameters can be derived by exploiting the homogeneity restrictions. Others can be estimated from a second-stage regression based on the residuals from equation (5.2.36), using the parameter estimates obtained from equation (5.2.43) and estimates of u.

The generalization to panel data follows along lines similar to those discussed for the case in which only technical inefficiency appears. In the presence of both technical and allocative inefficiency the translog normalized variable profit frontier and its associated input variable profit share equations are modified from those appearing in equations (5.2.30) and (5.2.31) to become

$$\begin{aligned}
\ln\frac{\pi_{it}}{p_{it}} &= \beta_o + \sum_n \beta_n \ln\frac{w_{nit}^s}{p_{it}e^{-ui}} \\
&+ \frac{1}{2}\sum_n\sum_k \beta_{nk}\ln\frac{w_{nit}^s}{p_{it}e^{-ui}}\ln\frac{w_{kit}^s}{p_{it}e^{-ui}} + \sum_q \gamma_q \ln z_{qit} \\
&+ \frac{1}{2}\sum_q\sum_r \gamma_{qr}\ln z_{qit}\ln z_{rit} \\
&+ \sum_n\sum_q \gamma_{nq}\ln\frac{w_{nit}}{p_{it}e^{-ui}}\ln z_{qit} + \ln\left[1-\sum_n(\exp\{\xi_n\}-1)S_{nit}^s\right] \\
&+ v_{it} - u_i,
\end{aligned} \tag{5.2.44}$$

$$S_{nit} = \frac{S_{nit}^s \exp\{\xi_n\}}{[1-\Sigma_k(\exp\{\xi_k\}-1)S_{kit}^s]} + \eta_{nit}, \qquad n = 1, \ldots, N. \tag{5.2.45}$$

where $w_n^s = w_n \cdot \exp\{-\xi_n\}$ $\forall n$ and the S_n^s are defined in equation (5.2.35). As before, v_{it} and the η_{nit} are stochastic noise error components. The model in equations (5.2.44) and (5.2.45) is structurally similar to the model in equations (5.2.30) and (5.2.31), apart from the allocative inefficiency terms, and estimation proceeds as in the previous case. Details are provided in Kumbhakar (1996b).

We concluded Section 5.2.1 with the observation that incorporating quasi-fixed inputs into a primal approach to the estimation of variable profit efficiency, although theoretically desirable, raises no new econometric issues. The same conclusion applies to a dual approach. Since no new econometric issues are raised in either approach, we ignore quasi-fixed inputs and consider the structurally simpler problem of estimating profit efficiency in the next two sections.

5.3 MULTIPLE-OUTPUT MODELS

In this section we provide a brief overview of the estimation of technical, allocative, and profit inefficiencies when producers produce multiple outputs and seek to maximize profit. We continue to measure technical inefficiency with an output orientation, although the analysis of this and Section 5.2 can be modified to adopt an input orientation. Sections 5.3.1 and 5.3.2 are structured similarly to Sections 5.2.1 and 5.2.2, with the former taking a primal approach and the latter taking a dual approach. There are two differences between this section and Section 5.2. First, here we consider the estimation of profit efficiency rather than of variable profit efficiency. The second difference concerns the modeling and estimation of technical and allocative inefficiencies, and of their impact on profit, in the presence of multiple outputs.

5.3.1 The Primal Distance Function Approach

Let $y = (y_1, \ldots, y_M) \geq 0$ be a vector of M outputs marketed at prices $p = (p_1, \ldots, p_M) > 0$, and let $x = (x_1, \ldots, x_N) \geq 0$ be a vector of N inputs purchased at prices $w = (w_1, \ldots, w_N) > 0$. The structure of production technology is characterized by the output distance function $D_o(x, y; \beta) \leq 1$, where β is a vector of technology parameters to be estimated. Output-oriented technical inefficiency is introduced by writing $D_o(x, y; \beta) = e^{-u} \leq 1 \Leftrightarrow D_o(x, ye^u; \beta) = 1$ for $u \geq 0$, since $ye^u \geq y$ is a technically efficient output vector and $D_o(x, y; \beta)$ is linearly homogeneous in y. If we allow for technical inefficiency but maintain the assumption of allocative efficiency, the producer's profit maximization problem can be written as

$$\max_{y,x}\{p^T y - w^T x : D_o(x, ye^u;\beta) = 1\}, \tag{5.3.1}$$

the first-order conditions for which can be rearranged to yield

$$\frac{w_n}{p_m e^{-u}} = -\frac{\partial D_o(x, ye^u;\beta)/\partial x_n}{\partial D_o(x, ye^u;\beta)/\partial y_m e^u}. \tag{5.3.2}$$

Equation (5.3.2), together with the distance function $D_o(x, ye^u; \beta) = 1$ in equation (5.3.1), can in principle be solved for $y_m(pe^{-u}, w; \beta)$ $\forall m$ and $x_n(pe^{-u}, w; \beta)$ $\forall n$. These are the output supply equations and input demand equations that maximize profit in the presence of technical inefficiency and that are to be estimated in order to obtain estimates of technical inefficiency and its impact on profit. Unfortunately "in principle" rarely occurs in practice. We now turn to one case in which it does, and then we consider the options available when it does not.

CET/Cobb–Douglas If the output distance function is separable, it can be written as $D_o(x, ye^u; \beta) = g(ye^u; \beta) - f(x; \beta)$. If, in addition, the functions $g(ye^u; \beta)$ and $f(x; \beta)$ are analytically tractable, an estimable model results. One such specification is the CET/Cobb–Douglas model of Powell and Gruen (1968), in which

$$g(ye^u;\beta) = \left\{\sum_m \delta_m (y_m e^u)^c\right\}^{1/c} = \left\{\sum_m \delta_m (y_m)^c\right\}^{1/c} e^u,$$

$$f(x;\beta) = \prod_n x_n^{\beta_n}, \tag{5.3.3}$$

where $c > 1$, $\delta_m > 0$ $\forall m$, $\Sigma_m\delta_m = 1$, $\beta_n > 0$ $\forall n$, and $\Sigma_n\beta_n = r < 1$. The partial elasticity of substitution between each pair of outputs is the constant $c/(1 - c) < 0$.

A lot of tedious algebra generates the profit-maximizing output supply and input demand equations

$$y_m(pe^{-u}, w;\beta) = I(p)^{(cr-1)/c(1-r)} \cdot p_m^{1/(c-1)} \cdot \delta_m^{-1/(c-1)} \prod_n \left(\frac{\alpha_n}{w_n}\right)^{\alpha_n/(1-r)} \cdot e^{-u/1-r}, \tag{5.3.4}$$

$$x_n(pe^{-u}, w;\beta) = I(p)^{(c-1)/c(1-r)} \cdot \prod_k \left(\frac{\alpha_k}{w_k}\right)^{[\alpha_k c+(1-rc)\delta_{nk}]/(1-rc)} \cdot e^{-u/(1-r)}, \tag{5.3.5}$$

where $I(p) = \Sigma_m \delta_m^{-c/(c-1)} p_m^{c/(c-1)}$, and $\delta_{nk} = 1$ if $n = k$ and $\delta_{nk} = 0$ if $n \neq k$. Equations (5.3.4) and (5.3.5) can be used to derive the expression for maximized profit, conditional on the presence of technical inefficiency,

$$\begin{aligned}\pi = \pi(pe^{-u}, w;\beta) &= \sum_m p_m y_m(pe^{-u}, w;\beta) - \sum_n w_n x_n(pe^{-u}, w;\beta)\\ &= (1-r)\cdot I(p)^{(c-1)/c(1-r)} \cdot \prod_n \left(\frac{\beta_n}{w_n}\right)^{\beta_n(1-r)} \cdot \exp\left\{\frac{-u}{1-r}\right\}\\ &= \pi(p, w;\beta)\cdot g(u). \end{aligned} \tag{5.3.6}$$

It follows that technical inefficiency reduces profit by $100[1/(1 - r)u]\%$. The fact that $\pi(pe^{-u}, w; \beta) = \pi(p, w; \beta)\cdot g(u)$ follows from the homogeneity of the CET/Cobb–Douglas technology.

Estimation of the system of equations (5.3.4) and (5.3.5) is straightforward. We append classical error terms η_j, $j = 1, \ldots, M + N$, to the equations and make the distributional assumptions:

(i) $\eta = (\eta_1, \ldots, \eta_{M+N})' \sim \text{iid } N(0, \Sigma)$.

(ii) $u \sim \text{iid } N^+(0, \sigma_u^2)$.

(iii) The elements of η are independent of u.

With these assumptions the log likelihood function is similar to that for equations (5.2.24). Defining $\Xi_i = \gamma u_i + \eta_i$, where $\gamma = -\iota 1/(1 - r)$ and ι is an $(M + N) \times 1$ vector of ones, the density function of Ξ_i can be written as

$$f(\Xi_i) = \frac{2\sigma\exp\{-b_i/2\}}{(2\pi)^{(M+N)/2}|\Sigma|^{1/2}\sigma_u}\Phi(\Xi_i'\Sigma^{-1}\gamma\sigma), \tag{5.3.7}$$

where $\sigma^2 = (1/\sigma_u^2 + \gamma'\Sigma^{-1}\gamma)^{-1}$ and $b_i = \Xi_i'\Sigma^{-1}\Xi_i - \sigma^2(\Xi_i'\Sigma^{-1}\gamma)^2$. The log likelihood function is

$$\ln L = \text{constant} - \frac{I}{2}\ln|\Sigma| + I\ln\sigma + \sum_i \ln\Phi(\Xi_i'\Sigma^{-1}\gamma\sigma)$$
$$- I\ln\sigma_u - \frac{1}{2}\sum_i b_i. \tag{5.3.8}$$

Once the parameters have been estimated, estimates of producer-specific technical efficiencies can be obtained from equations (5.2.28) and (5.2.29), in which b is replaced by γ.

With panel data and an assumption of time-invariant technical inefficiencies, the term $[-1/(1-r)]u$ in equations (5.3.4) and (5.3.5) is replaced with $\Sigma_i d_i D_i$, where the D_i are producer dummy variables. The system can be estimated using the nonlinear ITSUR technique, after which technical inefficiencies can be estimated relative to that of the most efficient producer as before.

Translog If $D_o(x, ye^u; \beta)$ takes the translog form, then it follows from Section 3.2.3 that it can be written as

$$-\ln|y| = \beta_o + \sum_n \beta_n \ln x_n + \sum_m \beta_m \ln\frac{y_m}{|y|} + \frac{1}{2}\sum_n\sum_k \beta_{nk}\ln x_n \ln x_k$$
$$+\frac{1}{2}\sum_m\sum_j \beta_{mj}\ln\frac{y_m}{|y|}\ln\frac{y_j}{|y|} + \sum_n\sum_m \beta_{nm}\ln x_n \ln\frac{y_m}{|y|} + u. \tag{5.3.9}$$

It is not possible to derive expressions for the profit-maximizing output supply and input demand equations, and the resulting profit equation, associated with the translog output distance function, so the approach adopted within the CET/Cobb–Douglas framework is infeasible. Even if the separability restrictions $\beta_{nm} = 0\ \forall n, m$ are imposed on the translog output distance function, derivation of profit-maximizing output supply and input demand equations, and the resulting profit equation, would be infeasible. This leaves two options, both of which suffer from the same problem. The first option is to estimate equation (5.3.9) by itself, after adding a random-noise error component. Although the ALS MLE technique can be applied, after which technical inefficiencies can be predicted using the JLMS decomposition technique, the parameter estimates will be biased and inconsistent because the regressors are not exogenous within a profit maximization framework. The second option is to incorporate additional equations based on the derivative property $\partial D_o(x, y; \beta)/\partial y_m = p_m^s/R^s\ \forall m$, where the p_m^s are output shadow prices and $R^s = \Sigma_m p_m^s y_m$ is shadow revenue. Although this will improve statistical efficiency, it does not address the endogeneity problem. We refer the reader to Atkinson and Primont (1998), who have developed an input-oriented version of such a system, and have proposed the use of instrumental variables estimation techniques to deal with the endogeneity problem.

5.3.2 The Dual Variable Profit Frontier Approach

The material in this section is a modest extension of the material in Section 5.2.2. We begin by revisiting the translog specification, and we continue by examining another flexible functional form.

Translog Allowing for technical inefficiency but maintaining the assumption of allocative efficiency, the translog profit frontier can be written as

$$\ln\pi = \ln\pi(pe^{-u}, w; \beta) = \beta_o + \sum_m \beta_m \ln(p_m e^{-u}) + \sum_n \beta_n \ln w_n$$
$$+\frac{1}{2}\sum_m\sum_j \beta_{mj}\ln(p_m e^{-u})\ln(p_j e^{-u}) + \frac{1}{2}\sum_n\sum_k \beta_{nk}\ln w_n \ln w_k$$
$$+\sum_m\sum_n \beta_{mn}\ln(p_m e^{-u})\ln w_n. \tag{5.3.10}$$

where the linear homogeneity restrictions are now $\Sigma_m\beta_m + \Sigma_n\beta_n = 1$, $\Sigma_m\beta_{jm} + \Sigma_m\beta_{jn} = 0\ \forall j$, and $\Sigma_n\beta_{km} + \Sigma_n\beta_{kn} = 0\ \forall k$. As usual, these restrictions can be imposed by writing the profit frontier in normalized form. The associated output profit share equations and input profit share equations can be derived using Hotelling's lemma and written as

$$R_m = \frac{p_m y_m}{\pi} = \beta_m + \sum_j \beta_{mj}\ln(p_j e^{-u}) + \sum_n \beta_{mn}\ln w_n,$$
$$m = 1, \ldots, M, \tag{5.3.11}$$

$$S_n = \frac{w_n x_n}{\pi} = -\left(\beta_n + \sum_m \beta_{mn}\ln(p_m e^{-u}) + \sum_k \beta_{nk}\ln w_k\right),$$
$$n = 1, \ldots, N. \tag{5.3.12}$$

As in the single-output case it is possible to express the profit function in equation (5.3.10) as $\ln \pi = \ln \pi(p, w; \beta) + \ln \pi_u$, where $\ln \pi_u$ represents the percentage profit loss due to technical inefficiency and is given by

$$\ln \pi_u = -u\left[\sum_m \beta_m + \sum_m \sum_j \beta_{mj} \ln p_j + \sum_m \sum_n \beta_{mn} \ln w_m\right] + \frac{u^2}{2} \sum_m \sum_j \beta_{mj}. \tag{5.3.13}$$

Thus even in the multiple-output case there is a closed-form solution for the profit loss due to technical inefficiency. Since $\ln \pi_u$ depends on all prices, omitting technical inefficiency or incorporating its impact in a simple additive form leads to biased and inconsistent parameter estimates.

In the absence of inefficiency, estimation would normally be based on the profit frontier with $(M + N - 1)$ profit share equations added to improve statistical efficiency. However in the present case technical inefficiency interacts with prices in the profit frontier, but not (after a simple rearrangement of terms so as to isolate u) in the profit share equations. Hence we base estimation on just the profit share equations. After appending classical error terms, the rearranged profit share equations can be written as

$$R_m = \beta_m + \sum_j \beta_{mj} \ln p_j + \sum_n \beta_{mn} \ln w_n - \beta_{mo} u + \eta_m, \quad m = 1, \ldots, M, \tag{5.3.14}$$

$$S_n = -\left(\beta_n + \sum_m \beta_{mn} \ln p_m + \sum_k \beta_{nk} \ln w_k\right) + \beta_{no} u + \eta_n, \quad n = 1, \ldots, N, \tag{5.3.15}$$

where $\beta_{mo} = \Sigma_j \beta_{mj}\ \forall m$ and $\beta_{no} = \Sigma_m \beta_{mn}\ \forall n$. A system of $(M + N - 1)$ of these share equations can be estimated, in cross-sectional or panel data contexts, exactly as in the single-output case discussed under Translog I in Section 5.2.2.

The model just discussed can be generalized to allow for allocative inefficiencies, just as the single-output **Translog I** model was generalized to create **Translog II** in Section 5.2.2. All that is required is to replace observed prices (p, w) with shadow prices $(p^s, w^s) = (p_1, p_2\exp\{-\xi_{21}\}, \ldots, p_M\exp\{-\xi_{M1}\}, w_1\exp\{-\zeta_{11}\}, \ldots, w_N\exp\{-\zeta_{N1}\})$ and proceed as in **Translog II**. Modeling procedures and estimation strategies are discussed in Kumbhakar (1996a).

Normalized Quadratic An alternative flexible functional form, the normalized quadratic, has also been used to model and estimate profit efficiency in the presence of multiple outputs and both technical and allocative inefficiency. Such a model, due originally to Fuss (1977) and McFadden (1978) and generalized by Diewert and Wales (1987) and Diewert and Ostensoe (1988), has some interesting features. In the formulation of Akhavein, Berger, and Humphrey (1997), the normalized quadratic profit frontier is written as

$$\frac{\pi}{q_1} = \sum_{j>1} (\beta_j - \tau_j)\left(\frac{q_j}{q_1}\right) + \sum_{j>1} \sum_{k>1} \beta_{jk} \left(\frac{1-\xi_j}{2}\right) \xi_k \left(\frac{q_j q_k}{q_1^2}\right) - \tau_1, \tag{5.3.16}$$

where $q = (p', w')'$ is an $(M + N)$-dimensional vector of netput prices and $q_1 = p_1$. The parameters β_j and $\beta_{jk} = \beta_{kj}$ are technology parameters to be estimated. Hotelling's lemma generates the system of netput equations

$$z_j = \beta_j + \sum_{k>1} \beta_{jk} \xi_k \left(\frac{q_k}{q_1}\right) - \tau_j, \qquad j = 2, \ldots, M+N, \tag{5.3.17}$$

where $z = (y, -x)$ is an $(M + N)$-dimensional netput quantity vector.

Inefficiencies are represented by the τ and ξ terms. Allocative inefficiencies are modeled as usual, by replacing actual price ratios (q_j/q_1) with shadow price ratios $\xi_j(q_j/q_1)$. The profit loss associated with allocative inefficiencies is calculated as

$$\pi(q; \beta, \tau, 1) - \pi(q; \beta, \tau, \xi) = \sum_{j>1} \sum_{k>1} \beta_{jk} \left[\frac{1}{2} - \left(\frac{1-\xi_j}{2}\right) \xi_k\right] \left(\frac{q_j q_k}{q_1}\right). \tag{5.3.18}$$

Equation (5.3.18) can itself be decomposed into profit loss attributable to (i) input mix inefficiencies, (ii) output mix inefficiencies, and (iii) input/output mix inefficiencies. Technical inefficiency is modeled through the τ vector, with netput z_j being τ_j units beneath the value that would maximize profit in the face of shadow prices. The profit loss associated with technical inefficiency is calculated as

$$\pi(q;\beta,\tau,\xi)-\pi(q;\beta,0,\xi)=\sum_j \tau_j q_j. \tag{5.3.19}$$

The generalized quadratic specification differs from the translog specification in a number of ways. First, it is not logarithmic, and so allows for nonpositive values of profit and, in its more popular normalized variable profit form, for nonpositive values of quasi-fixed netputs as well. Second, convexity can be imposed, without reducing flexibility, by constraining the symmetric $(M+N)\times(M+N)$ matrix $\beta=[\beta_{jk}]$ to be positive semidefinite by replacing β with bb', where b is a lower triangular matrix with zero elements in its first column. Third, technical and allocative inefficiencies do not interact, making it possible to calculate the profit loss associated with either one independently of the other. Finally, and unfortunately, technical inefficiency is measured nonradially, as in the Lovell and Sickles (1983) generalized Leontief model, with each of the $M+N$ technical efficiency measures being dependent on the units in which its corresponding netput is measured. Nonetheless, profit loss associated with technical inefficiency is independent of units of measurement.

A subset of $M+N$ independent equations from the system (5.3.16) and (5.3.17) can be estimated, in either a cross-sectional or a panel data context, by adding classical error terms and using nonlinear ITSUR. In a cross-sectional context, and also in a panel data context when the number of producers is large, it is necessary to treat the allocative inefficiency parameters ξ_j as varying across netput ratios but fixed across producers. The composed error terms (involving technical inefficiencies and random noise) can be treated within a conventional fixed-effects framework or, what is essentially the same thing, within a "distribution-free" framework discussed in Section 4.4.2.

5.4 ALTERNATIVE PROFIT FRONTIERS

Underlying the profit frontier $\pi(p,w;\beta)$ [and the variable profit frontier $\pi(p,w,z;\beta)$ as well, z being a vector of quasi-fixed inputs] is the assumption that prices are exogenous and that producers seek to maximize profit (or variable profit) by selecting outputs and inputs under their control. One justification for exogeneity of prices is that producers operate in competitive markets. Suppose, to the contrary, that producers have some degree of monopoly power in their product markets. Under monopoly, the demands would be exploited to determine output prices and quantities jointly, and only input prices would be exogenous. In this context neither a traditional cost frontier (which treats outputs as being exogenous) nor a traditional profit frontier (which treats output prices as being exogenous) would provide an appropriate framework within which to evaluate producer performance.

Recently Humphrey and Pulley (1997) and others have introduced the notion of an "alternative" profit frontier to bridge the gap between a cost frontier and a profit frontier. An alternative profit frontier is defined as

$$\pi^A(y,w;\beta,\delta)=\max_{p,x}\{p^Ty-w^Tx: g(p,y,w;\delta)=0,\ D_o(x,y;\beta)\le 1\}, \tag{5.3.20}$$

where the endogenous variables are (p,x) and the exogenous variables are (y,w), $D_o(x,y;\beta)$ is the output distance function characterizing the structure of production technology, and $g(p,y,w;\delta)$ represents what Humphrey and Pulley refer to as the producer's "pricing opportunity set," which captures the producer's ability to transform exogenous (y,w) into endogenous product prices p. Thus $\pi^A(y,w;\beta,\delta)$ has the same dependent variable as the standard profit frontier and the same independent variables as the standard cost frontier. However $\pi^A(y,w;\beta,\delta)$ is not dual to $D_o(x,y;\beta)$, because it incorporates both the structure of production technology (incorporated in the parameter vector β) and the structure of the pricing opportunity set (incorporated in the parameter vector δ). Moreover, without specifying the properties satisfied by the function $g(p,y,w;\delta)$, it is not possible to specify the properties satisfied by $\pi^A(y,w;\beta,\delta)$, although it is reasonable to assume that $\pi^A(y,w;\beta,\delta)$ is nondecreasing in the elements of y and nonincreasing in the elements of w.

In the absence of Shephard–Hotelling derivative properties, it is not possible to specify a system of equations on which estimation can be based, as would be the case with standard cost or profit frontiers. Consequently the alternative profit frontier must be estimated as a single-equation model, once a functional form is assigned to $\pi^A(y,w;\beta,\delta)$ and an assumption is made concerning error structure.

One possibility is to assume that $\pi^A(y, w; \beta, \delta)$ takes a translog functional form with composed error structure $(v - u)$, with $v \sim$ iid $N(0, \sigma_v^2)$ and $u \sim$ iid $N^+(0, \sigma_u^2)$, as in the conventional stochastic frontier framework.

5.5 A GUIDE TO THE LITERATURE

The theory underlying profit functions/frontiers and variable profit functions/frontiers goes back to Hotelling (1932). Modern treatments are available in Lau (1972, 1976, 1978), Diewert (1973), and McFadden (1978). The result of Lau (1978) showing the implication of homogeneity of a single-output technology for the structure of the dual variable profit frontier, which is utilized in Section 5.2, also carries over to multiple-output technologies in Section 5.3, in which case homogeneity is replaced with almost homogeneity. Lau also shows the implication of separability of the multiple-output distance function (which he calls a transformation function) for the dual profit frontier.

The use of shadow prices to model allocative inefficiencies within a profit-seeking context is due to Lau and Yotopoulos (1971) and Yotopoulos and Lau (1973), although they assumed a restrictive Cobb–Douglas functional form, as did subsequent research contained in *Food Research Studies* (1979). The generalization to flexible functional forms within a profit-seeking context is apparently due to Atkinson and Halvorsen (1980) (translog, assuming technical efficiency) and Lovell and Sickles (1983) (generalized Leontief, allowing for both technical and allocative inefficiency).

What we have referred to as the primal approach to the estimation of technical, allocative, and profit efficiency has been used by Kumbhakar (1987) (Cobb–Douglas), Kumbhakar, Biswas, and Bailey (1989) (Cobb–Douglas), Kalirajan (1990) (translog), Seale (1990) (Cobb–Douglas), Kalirajan and Obwana (1994a, b) (Cobb–Douglas with random coefficients), and Kumbhakar (1994) (translog).

A detailed treatment of the primal approach in the multiple-output context, with and without separability imposed, is provided by Hasenkamp (1976).

Direct estimation of the dual profit or normalized profit frontier has been used by Hollas and Stansell (1988) (translog), Ali and Flinn (1989) (translog with random coefficients), Atkinson and Kerkvliet (1989) (translog), Kumbhakar and Bhattacharyya (1992) (translog), Berger, Hancock, and Humphrey (1993) (generalized quadratic), Bhattacharyya and Glover (1993) (generalized quadratic), Ali, Parikh, and Shah (1994) (translog), Kumbhakar (1996b) (translog), Akhavein, Berger, and Humphrey (1997) (generalized quadratic), and Berger and Mester (1997) (Fourier), among many others.

Alternative profit frontiers have been formulated and estimated by Berger, Cummins, and Weiss (1996), Hasan and Hunter (1996), Akhavein, Berger, and Humphrey (1997), Berger and Mester (1997), Humphrey and Pulley (1997), and Lozano Vivas (1997). Each has used a single-equation model, although a variety of functional forms have been specified, and a variety of estimation techniques has been employed. Each has estimated technical inefficiency, and its impact on alternative profit, under an assumption of allocative efficiency.