8 The Estimation of Efficiency Change and Productivity Change

8.1 INTRODUCTION

Productivity change occurs when an index of outputs changes at a different rate than an index of inputs does. Two questions immediately arise. First, how can productivity change be measured? Second, what are the sources of measured productivity change? Diewert (1992) has answered the first question. Productivity change can be calculated using index number techniques to construct a superlative Fisher (1922) or Törnqvist (1936) productivity index. Both indexes require quantity and price information, as well as assumptions concerning the structure of technology and the behavior of producers, but neither requires the estimation of anything. Productivity change can also be calculated using nonparametric techniques, or estimated using econometric techniques, to construct what has come to be known as a Malmquist (1953) productivity index. These latter techniques do not require price information or technological and behavioral assumptions, but they do require calculation or estimation of a representation of production technology. Thus information on the structure of the technology that generates the quantity data can serve as a substitute for price data and assumptions.

A disadvantage of index number techniques is that they do not provide an answer to the second question, whereas nonparametric techniques and econometric techniques do. Although nonparametric techniques and econometric techniques are capable of providing answers to both questions, only the econometric approach is capable of doing so in a stochastic environment. In this chapter we show how to use econometric techniques to estimate the magnitude of productivity change, and then to decompose estimated productivity change into its various sources.

Early econometric studies of productivity change used a primal approach, based on the assumption of a single output and the estimation of a production function. More recently, with the development of duality theory, the majority of econometric studies of productivity change have used a dual approach, which allows for multiple outputs and which can be based on the estimation of either a cost function or a profit function. The estimation and decomposition of productivity change is by now well developed within a cost function framework, although it is much less well developed within a profit function framework.

Until very recently econometric models of productivity change ignored the contribution of efficiency change to productivity change; hence the word *function* in place of the word *frontier* in the preceding paragraph. Productivity change was initially allocated exclusively to shifts in production technology (the magnitude of neutral technical change); eventually roles were also assigned to the biases of technical change and the structure of the technology (scale economies). However if inefficiency exists, then efficiency change provides an independent contribution to productivity change. If efficiency change is omitted from the analysis, its omission leads to an overstatement of the unexplained residual, which Abramovitz (1956) aptly referred to as a "measure of our ignorance about the causes of economic growth," and also to an erroneous allocation of productivity change to its included sources. Accordingly, it is desirable to incorporate the possibility of efficiency change, both technical and allocative, into econometric models of productivity change. This is our objective in this chapter.

We begin by using a production frontier approach to obtain estimates of productivity change, and to decompose estimated productivity change into a technical change component, a returns to scale component, and a component associated with change in technical efficiency. We continue by doing essentially the same thing using a cost frontier approach, although we also include a component attributable to change in input allocative efficiency. Finally we use a profit frontier approach, and we demonstrate that the profit frontier approach gives essentially the same productivity change decomposition as the production frontier approach does, provided that the first-order conditions for profit maximization are used in the production frontier approach. A considerable advantage of the two dual approaches is that the productivity change decomposition can be conducted within a multiple-output framework.

We follow the format of preceding chapters and break down our discussion of the contribution of efficiency change to productivity change into three sections. In Section 8.2 we develop the primal (production frontier) approach to the estimation and decomposition of productivity change in the presence of technical inefficiency. The material in this section builds on the material developed in Chapter 3. In Sections 8.3 and 8.4 we develop a pair of dual (cost frontier and profit frontier) approaches to the estimation and decomposition of productivity change in the presence of both technical and allocative inefficiency. The material in these sections builds on the material developed in Chapters 4–6. Since the primary issue is one of productivity change, we assume throughout this chapter that panel data are available. Section 8.5 concludes with a guide to the relevant literature.

8.2 THE PRIMAL (PRODUCTION FRONTIER) APPROACH

In this section we develop a quantity-based approach to the estimation and decomposition of productivity change. We associate productivity change with either four or three components, depending on whether or not price information is also available. In Section 8.2.1 we develop an analytical framework in which productivity change is defined and the two decompositions are derived analytically. In the likely event that price information is unavailable, productivity change is decomposed into technical change, technical efficiency change, and the contribution of returns to scale. In Section 8.2.2 we discuss estimation of a composed error translog production frontier, in which the technical inefficiency error component is allowed to be time varying. We also show how the parameter estimates can be used to obtain estimates of the three components of productivity change.

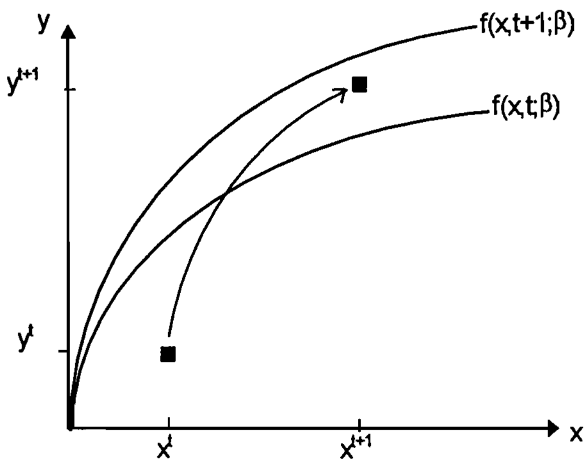

Figure 8.1 The Primal Approach to the Estimation and Decomposition of Productivity Change

The general structure of the primal approach is illustrated in Figure 8.1, in which a single input is used to produce a single output, and a producer expands from (x^t, y^t) to (x^{t+1}, y^{t+1}). Production technology is characterized by decreasing returns to scale, and technical progress has occurred between periods t and $t+1$, since $f(x, t+1; \beta) > f(x, t; \beta)$. Assuming away noise for the moment, it is clear that production is technically inefficient in both periods, since $y^t < f(x^t, t; \beta)$ and $y^{t+1} < f(x^{t+1}, t+1; \beta)$, and that technical efficiency has improved from period t to period $t+1$, since $[y^t/f(x^t, t; \beta)] < [y^{t+1}/f(x^{t+1}, t+1; \beta)]$. It is also clear that productivity growth has occurred, since $(y^{t+1}/x^{t+1}) > (y^t/x^t)$. The initial econometric problem is to attribute output growth to input growth and productivity growth. The estimated rate of productivity growth must then be decomposed into contributions associated with returns to scale, technical change, and change in technical efficiency.

8.2.1 The Analytical Framework

We start with the deterministic production frontier

$$y = f(x, t; \beta) \cdot \exp\{-u\}, \tag{8.2.1}$$

where y is the scalar output of a producer, $f(x, t; \beta)$ is the deterministic kernel of a stochastic production frontier with technology para-

meter vector β to be estimated, $x = (x_1, \ldots, x_N) \geq 0$ is an input vector, t is a time trend serving as a proxy for technical change, and $u \geq 0$ represents output-oriented technical inefficiency. Technical change is not restricted to be neutral with respect to the inputs; neutrality requires that $f(x, t; \beta) = A(t) \cdot g(x; \beta)$.

A primal measure of the rate of technical change is provided by

$$T\Delta = \frac{\partial \ln f(x, t; \beta)}{\partial t}. \tag{8.2.2}$$

$T\Delta \gtreqless 0$ according as technical change shifts the production frontier up, leaves it unchanged, or shifts it down. A primal measure of the rate of change in technical efficiency is given by

$$TE\Delta = -\frac{\partial u}{\partial t}. \tag{8.2.3}$$

$TE\Delta \gtreqless 0$ according as technical inefficiency declines, remains unchanged, or increases through time. $TE\Delta$ can be interpreted as the rate at which a producer moves toward or away from the production frontier, which itself may be shifting through time.

In the scalar output case a conventional Divisia index of productivity change is defined as the difference between the rate of change of output and the rate of change of an input quantity index, and so

$$\begin{aligned} T\dot{F}P &= \dot{y} - \dot{X} \\ &= \dot{y} - \sum_n S_n \dot{x}_n, \end{aligned} \tag{8.2.4}$$

where a dot over a variable indicates its rate of change [e.g., $\dot{y} = (1/y)(dy/dt) = d\ln y/dt$], $S_n = w_n x_n / E$ is the observed expenditure share of input x_n, $E = \Sigma_n w_n x_n$ is total expenditure, and $w = (w_1, \ldots, w_N) > 0$ is an input price vector. Totally differentiating equation (8.2.1) and inserting the resulting expression for $\dot{y}$ into equation (8.2.4) yields

$$\begin{aligned} T\dot{F}P &= T\Delta + \sum_n (\varepsilon_n - S_n)\dot{x}_n + TE\Delta \\ &= T\Delta + (\varepsilon - 1) \cdot \sum_n \left(\frac{\varepsilon_n}{\varepsilon}\right)\dot{x}_n + \sum_n \left[\left(\frac{\varepsilon_n}{\varepsilon}\right) - S_n\right]\dot{x}_n + TE\Delta, \end{aligned} \tag{8.2.5}$$

where $\varepsilon_n = \varepsilon_n(x, t; \beta) = x_n f_n(x, t; \beta)/f(x, t; \beta)$, $n = 1, \ldots, N$, are elasticities of output with respect to each of the inputs. The scale elasticity

$\varepsilon = \varepsilon(x, t; \beta) = \Sigma_n \varepsilon_n(x, t; \beta) \gtreqless 1$ provides a primal measure of returns to scale characterizing the production frontier. The relationship in equation (8.2.5) decomposes productivity change into a technical change component $[T\Delta]$, a scale component $[(\varepsilon - 1) \cdot \Sigma_n(\varepsilon_n/\varepsilon)\dot{x}_n]$, a technical efficiency change component $[TE\Delta]$, and an allocative inefficiency component $[\Sigma_n[(\varepsilon_n/\varepsilon) - S_n]\dot{x}_n]$. This decomposition of productivity change is very similar to the decomposition obtained by Bauer (1990a; 289).

Interpretation of three components is straightforward. If either production technology or technical efficiency is time invariant, then it makes no contribution to productivity change. The contribution of scale economies depends on technology and data. Under constant returns to scale input growth or contraction makes no contribution to productivity change. Nonconstant returns to scale makes a positive contribution to productivity change if the scale elasticity $\varepsilon(x, t; \beta) > 1$ and input use expands $[\Sigma_n(\varepsilon_n/\varepsilon)\dot{x}_n > 0]$, or if the scale elasticity $\varepsilon(x, t; \beta) < 1$ and input use contracts $[\Sigma_n(\varepsilon_n/\varepsilon)\dot{x}_n < 0]$.

Interpretation of the allocative inefficiency component is a bit more involved. This component clearly captures the impact of deviations of inputs' normalized output elasticities from their expenditure shares or, somewhat less clearly, of input prices from the value of their marginal products $[w_n \gtreqless pf_n(x, t; \beta) \cdot \exp\{-u\}, n = 1, \ldots, N]$, where p is the price at which y is sold. The allocative inefficiency component can represent either input allocative inefficiency $[f_n(x, t; \beta)/f_k(x, t; \beta) \neq w_n/w_k]$ or scale inefficiency $[f_n(x, t; \beta)/f_k(x, t; \beta) = w_n/w_k$ but $w_n \neq pf_n(x, t; \beta) \cdot \exp\{-u\}]$ or a combination of the two. If producers are allocatively efficient, then $f_n(x, t; \beta) = (w_n/p) \cdot \exp\{u\}$, $n = 1, \ldots, N$, the allocative inefficiency component vanishes, and the scale component becomes $(\varepsilon - 1) \cdot \Sigma_n S_n \dot{x}_n$.

If price information is unavailable the allocative inefficiency component cannot be calculated empirically, whether or not allocative inefficiency exists. In this case it is implicitly assumed that $S_n = (\varepsilon_n/\varepsilon)$ $\forall n$, and the decomposition in equation (8.2.5) simplifies to

$$T\dot{F}P = T\Delta + (\varepsilon - 1) \cdot \sum_n \left(\frac{\varepsilon_n}{\varepsilon}\right) \dot{x}_n + TE\Delta, \tag{8.2.6}$$

which contains only quantity information. If technical efficiency is time invariant, then the third component on the right-hand side of equation (8.2.6) drops out, and the decomposition collapses to that of Denny, Fuss, and Waverman (1981; 193), in which productivity change is composed of technical change and a scale economies effect. If technical efficiency is time invariant and constant returns to scale prevail, the second and third components on the right-hand side of equation (8.2.6) both drop out, and productivity change consists solely of technical change. Thus only in the presence of time-invariant technical efficiency, persistent allocative efficiency, and constant returns to scale is it possible to associate productivity change with technical change, as was customary in early econometric studies of productivity change.

8.2.2 Estimation and Decomposition

The components of productivity change derived in Section 8.2.1 can be estimated within a stochastic production frontier framework. We assume that we have panel data on I producers through T time periods, and that the time-varying production frontier can be expressed in translog form as

$$\ln y_{it} = \beta_o + \sum_n \beta_n \ln x_{nit} + \beta_t t + \frac{1}{2}\sum_n \sum_k \beta_{nk} \ln x_{nit} \ln x_{kit} + \frac{1}{2}\beta_{tt} t^2 + \sum_n \beta_{nt} \ln x_{nit} t + v_{it} - u_{it}, \tag{8.2.7}$$

where $v_{it} \sim$ iid $N(0, \sigma_v^2)$ is the random-noise error component and $u_{it} \geq 0$ is the technical inefficiency error component. Technical change is neutral with respect to inputs if, and only if, $\beta_{nt} = 0$ $\forall n$, and absent if, and only if, $\beta_t = \beta_{tt} = \beta_{nt} = 0$ $\forall n$. Since we assume that we do not have input price information, the productivity change decomposition we wish to implement is that appearing in equation (8.2.6).

Equation (8.2.7) is structurally similar to the panel data models considered in Section 3.3, and all of the estimation techniques considered there are applicable in this context. The only difference here is the appearance of t as a regressor intended to capture the effects of technical change. The principal econometric problem is to disentangle the two roles played by t: as a proxy for technical change in the deterministic kernel of the stochastic production frontier and as an indicator of technical efficiency change in the second error component.

In Chapter 3 we discussed several time-varying specifications for the technical inefficiency error component. These are: (i) $u_{it} = u_i \cdot \gamma(t)$, where $\gamma(t)$ is a parametric function of time and u_i is a nonnegative random variable [Kumbhakar (1990) and Battese and Coelli (1992)]; (ii) $u_{it} = u_i \cdot \gamma_t$, where the γ_t are time effects represented by time dummies and the u_i can be either fixed or random producer-specific effects [Lee and Schmidt (1993)]; and (iii) $u_{it} = \Omega_{1i} + \Omega_{2i}t + \Omega_{3i}t^2$, where the Ωs are producer-specific parameters [Cornwell, Schmidt, and Sickles (1990)].

In specification (i) the separate time effects can be disentangled by positing a nonlinear specification for $\gamma(t)$. We consider this model, with distributional assumptions $v_{it} \sim$ iid $N(0, \sigma_v^2)$ and $u_{it} = u_i \cdot \exp\{-\gamma(t - T)\}$ with $u_i \sim$ iid $N^+(\mu, \sigma_u^2)$. Based on these assumptions, one can derive the probability density function of the composite error term $v_{it} - u_{it}$, and hence the log likelihood function for the model in equation (8.2.7). Once maximum likelihood estimates of the parameters in (8.2.7) are obtained, one can use the best linear unbiased predictor of technical efficiency. Formulation and estimation of this model, although without time as a regressor, was considered in Section 3.3.2, and additional details are available in Battese and Coelli (1992).

Since our interest centers on the estimation of productivity change and its sources given in equation (8.2.6), we need to obtain estimates of $T\Delta$, $TE\Delta$, $\varepsilon_n\ \forall n$, and ε. These can be derived from the parameter estimates obtained from equation (8.2.7) by means of

$$\hat{T\Delta} = \hat{\beta}_t + \hat{\beta}_{tt}t + \sum_n \hat{\beta}_{nt} \ln x_{nit},$$

$$T\hat{E}\Delta = \hat{u}_i \cdot \hat{\gamma} \cdot \exp\{-\hat{\gamma}(t - T)\},$$

$$\hat{\varepsilon}_n = \hat{\beta}_n + \sum_k \hat{\beta}_{nk} \ln x_{kit} + \hat{\beta}_{nt}t, \qquad n = 1, \ldots, N,$$

$$\hat{\varepsilon} = \sum_n \left(\hat{\beta}_n + \sum_k \hat{\beta}_{nk} \ln x_{kit} + \hat{\beta}_{nt}t \right).$$

Once these components are estimated, the productivity change index in equation (8.2.6) can be estimated and decomposed for each producer. Notice that all three components of equation (8.2.6) are time and producer specific unless certain parametric restrictions are satisfied. $T\Delta$ varies across producers unless it is neutral with respect to inputs ($\beta_{nt} = 0\ \forall n$), and $T\Delta$ varies through time unless $\beta_{tt} = \beta_{nt} = 0\ \forall n$. $TE\Delta$ varies across producers through its u_i component, and

$TE\Delta$ varies through time, with the same trend for each producer, unless $\gamma = 0$. Finally $[(\varepsilon - 1) \cdot \Sigma_n(\varepsilon_n/\varepsilon)\dot{x}_{nit}]$ varies across producers and through time unless technology takes the linearly homogeneous Cobb–Douglas form ($\Sigma_n\beta_n = 1$ and $\beta_{nk} = \beta_{nt} = 0\ \forall n, k$) or unless $\dot{x}_{nit} = 0\ \forall n, i, t$. Each of these parametric restrictions is testable.

8.3 A DUAL (COST FRONTIER) APPROACH

In this section we use quantity and price information to estimate and decompose productivity change. This dual approach is based on a cost frontier. In Section 8.3.1 we develop a pair of analytical frameworks in which dual productivity change is expressed in terms of technical change, technical efficiency change, and the contribution of scale economies, as in the primal approach given in equation (8.2.6). However the dual approach also enables us to derive a component that captures the effect of change in input allocative inefficiency. The two analytical frameworks differ in that one models the *costs* of technical and input allocative inefficiency as error components in a stochastic cost frontier, whereas the other models the *magnitudes* of technical and input allocative inefficiency by treating them as additional parameters to be estimated. The two frameworks generate identical technical change and scale components, but they generate different technical efficiency change and input allocative efficiency change components. In Section 8.3.2 we consider estimation and decomposition of productivity change using the second framework and a translog stochastic cost frontier. Sections 8.3.1 and 8.3.2 are consistent with Sections 8.2.1 and 8.2.2 in the sense that both assume that a single output is produced. Section 8.3.3 exploits the ability of a cost frontier to easily accommodate multiple outputs, and develops a decomposition of productivity change in a multiple-output context. This section is brief, because our emphasis is on introducing efficiency change into a productivity change decomposition rather than on expanding the dimensionality of output space.

The general structure of the dual cost frontier approach is illustrated in Figure 8.2, in which N inputs are used to produce a single output, and a producer expands from (y^t, E^t) to (y^{t+1}, E^{t+1}). Production technology is characterized by decreasing returns to scale, and

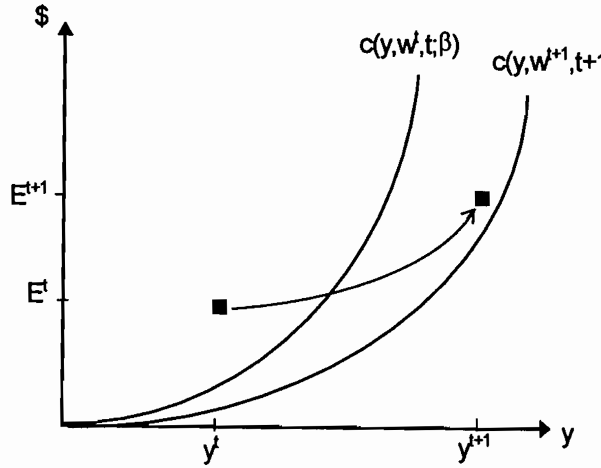

Figure 8.2 The Dual Approach to the Estimation and Decomposition of Productivity Change

technical progress has occurred between periods t and $t+1$, since $c(y, w^{t+1}, t+1; \beta) < c(y, w^t, t; \beta)$ and since we assume that $w^{t+1} \geq w^t$. Assuming away noise for the moment, it is clear that production is cost inefficient in both periods, since $E^t > c(y^t, w^t, t; \beta)$ and $E^{t+1} > c(y^{t+1}, w^{t+1}, t+1; \beta)$, and that cost efficiency has improved from period t to period $t+1$, since $[E^{t+1}/c(y^{t+1}, w^{t+1}, t+1; \beta)] < [E^t/c(y^t, w^t, t; \beta)]$. It is also clear that productivity growth has occurred, since $(E^{t+1}/y^{t+1}) < (E^t/y^t)$. The initial econometric problem is to attribute expenditure growth to output growth and productivity growth. The estimated rate of productivity growth must then be decomposed into contributions associated with returns to scale, technical change, and change in cost efficiency, which itself may be due to change in technical and/or input allocative efficiency.

8.3.1 The Analytical Framework

A dual cost minimization framework is widely used in the productivity literature to estimate and decompose productivity change. Although it is possible to use a primal framework and invoke the first-order conditions for cost minimization to derive an expression for productivity change, we prefer to use a dual framework. The main

reason for using a dual framework is that it is useful from an estimation point of view. We begin with a completely general formulation, which is not parameterized and which has been used as the basis for empirical research when the contribution of efficiency change to productivity change is ignored. We then continue with a preferred parameterized formulation on which estimation can be based in the presence of efficiency change.

Consider the deterministic cost frontier

$$\ln E = \ln c(y,w,t;\beta) + u_T + u_A, \tag{8.3.1}$$

where $E = w^T x$ is total expenditure, $x = (x_1, \ldots, x_N) \geq 0$ is an input vector, $w = (w_1, \ldots, w_N) > 0$ is an input price vector, $y \geq 0$ is scalar output, t is a time trend that serves as a proxy for technical change, $c(y, w, t; \beta)$ is the deterministic kernel of a stochastic cost frontier with technology parameter vector β to be estimated, $u_T \geq 0$ is the cost of technical inefficiency, and $u_A \geq 0$ is the cost of input allocative inefficiency. Totally differentiating equation (8.3.1), solving for $\dot{y}$, and substituting the expression for $\dot{y}$ into equation (8.2.4) yields an initial expression for dual productivity change

$$\begin{aligned} T\dot{F}P &= [1-\varepsilon(y,w,t;\beta)]\dot{y} - \dot{c}(y,w,t;\beta) - \sum_n S_n(y,w,t;\beta)\dot{w}_n \\ &\quad - \sum_n \left(\frac{w_n x_n}{E}\right)\dot{x}_n + \dot{E} - \frac{\partial u_T}{\partial t} - \frac{\partial u_A}{\partial t}, \end{aligned} \tag{8.3.2}$$

where the cost elasticity $\varepsilon(y, w, t; \beta) = \partial \ln c(y, w, t; \beta)/\partial \ln y = [\varepsilon(x, t; \beta)]^{-1}$ provides a dual measure of returns to scale and the $S_n(y, w, t; \beta) = \partial \ln c(y, w, t; \beta)/\partial \ln w_n$, $n = 1, \ldots, N$, are efficient input cost shares. Next, noting that $\dot{E} - \Sigma_n(w_n x_n/E)\dot{x}_n = \Sigma_n(w_n x_n/E)\dot{w}_n$ and substituting this equality into equation (8.3.2) yields the final expression for dual productivity change

$$\begin{aligned} T\dot{F}P &= [1-\varepsilon(y,w,t;\beta)]\dot{y} - \dot{c}(y,w,t;\beta) \\ &\quad + \sum_n [S_n - S_n(y,w,t;\beta)]\dot{w}_n - \frac{\partial u_T}{\partial t} - \frac{\partial u_A}{\partial t}, \end{aligned} \tag{8.3.3}$$

where the $S_n = w_n x_n/E$, $n = 1, \ldots, N$, are actual input cost shares.

Equation (8.3.3) provides a decomposition of dual productivity change that is identical to that obtained by Bauer (1990a; 291). The first component is a scale effect, which makes no contribution to productivity change if either $\varepsilon(y, w, t; \beta) = 1$ or $\dot{y} = 0$. However output growth in the presence of scale economies $[\varepsilon(y, w, t; \beta) < 1]$ contributes to productivity growth, as does output contraction in the presence of diseconomies of scale $[\varepsilon(y, w, t; \beta) > 1]$. Conversely output growth in the presence of diseconomies of scale retards productivity growth, as does output contraction in the presence of scale economies. The second component is a technical change effect that shifts the cost frontier down if technical change is progress or up if technical change is regress. The third component captures the impact of deviations of actual input cost shares from efficient input cost shares. This component is zero if the input mix is allocatively efficient, since then $S_n = S_n(y, w, t; \beta)\ \forall n$. However this term is also zero in the presence of input allocative inefficiency if all input prices change at the same rate because $\Sigma_n[S_n - S_n(y, w, t; \beta)] = 0$. The final two components capture the separate contributions to productivity change of changes in the costs of technical and input allocative inefficiency.

The dual productivity change decomposition given in equation (8.3.3) parallels the two primal productivity change decompositions given in equations (8.2.6) and (8.2.5), although not exactly. Both primal and dual decompositions contain components measuring the contributions of technical change and scale economies to productivity change. However the dual technical change component in equation (8.3.3) is not identical to the primal technical change component in equations (8.2.6) and (8.2.5) unless constant returns to scale prevail. Ohta (1974) has shown that

$$-\dot{c}(y,w,t;\beta) = \varepsilon(y,w,t;\beta)\cdot T\Delta, \tag{8.3.4}$$

so that $[-\dot{c}(y, w, t; \beta)] \gtreqless T\Delta$ according as $\varepsilon(y, w, t; \beta) = [\varepsilon(x, t; \beta)]^{-1} \gtreqless 1$. The dual scale component in equation (8.3.3) is related to the primal scale component in equations (8.2.5) and (8.2.6) by way of

$$[1-\varepsilon(y,w,t;\beta)]\dot{y} = (\varepsilon - 1)\cdot \sum_n \left(\frac{\varepsilon_n}{\varepsilon}\right)\dot{x}_n + [1-\varepsilon^{-1}]\cdot T\Delta. \tag{8.3.5}$$

Thus the two scale components are equal under constant returns to scale, in which case both are zero, or in the absence of technical change. In addition, the primal and dual technical efficiency change components differ, the former measuring the contribution of changes in the magnitude of output-oriented technical inefficiency and the latter measuring the contribution of changes in the cost of technical

inefficiency whose orientation is as yet unspecified. Finally the primal formulation in equation (8.2.5) contains one allocative efficiency change component, and the simpler primal formulation in equation (8.2.6) contains no allocative efficiency change component, whereas the dual formulation in equation (8.3.3) contains a pair of allocative efficiency change components.

The problem with Bauer's dual productivity change decomposition given in equation (8.3.3) lies with the two error components u_T and u_A. We saw in Chapter 4 how difficult it is to estimate the costs of technical and input allocative inefficiencies as error components within a flexible cost frontier model. In the present context this leaves us with two options. The first is to simplify equation (8.3.1) by replacing the two error components u_T and u_A with a single error component u_{T+A} representing the cost of both types of inefficiency. This leads to the replacement of the two terms $-\partial u_T/\partial t$ and $-\partial u_A/\partial t$ in equations (8.3.2) and (8.3.3) with the single term $-\partial u_{T+A}/\partial t$ representing change in the cost of both types of inefficiency. The second option is to base our decomposition of productivity change on a parametric model along the lines developed in Chapter 6. Since it is difficult enough to interpret, much less to estimate, the dual technical and allocative efficiency change components within a composed error framework, we follow the second option by developing a partially parameterized dual formulation in which the contributions of changes in technical and input allocative efficiency are relatively easy to interpret and estimate.

Consider the cost minimization problem

$$\min_x\{w^T x:\ y = f(x\cdot\exp\{-\tau\},t;\beta)\}, \tag{8.3.6}$$

where $f(x, t; \beta)$ is the production frontier and $\exp\{-\tau\}$ represents input-oriented technical inefficiency with $\tau \geq 0$. In addition to allowing for technical inefficiency, we also allow for input allocative inefficiency by expressing the first-order conditions for equation (8.3.6) as

$$\frac{f_n(x\cdot\exp\{-\tau\},t;\beta)}{f_1(x\cdot\exp\{-\tau\},t;\beta)} = \frac{w_n}{w_1}\cdot\exp\{-\xi_n\}, \qquad n = 2,\ldots,N. \tag{8.3.7}$$

This formulation allows the technically efficient input vector $x\cdot\exp\{-\tau\}$ to be allocatively inefficient relative to the observed input price vector w. A producer is allocatively efficient relative to w if, and only if, $\xi_n = 0$, $n = 2, \ldots, N$. The technically efficient input vector $x \cdot \exp\{-\tau\}$ is, however, allocatively efficient relative to the shadow input price vector $[w_1, w_2 \cdot \exp\{-\xi_2\}, \ldots, w_N \cdot \exp\{-\xi_N\}]$.

Defining the shadow input price vector as $w^s = [w_1, w_2 \cdot \exp\{-\xi_2\}, \ldots, w_N \cdot \exp\{-\xi_N\}]$ and the corresponding shadow cost frontier (which does not incorporate technical inefficiency) as $c^s(y, w^s, t; \beta) = \Sigma_n w_n^s x_n$, and substituting w^s into the optimization problem (8.3.6), generates the shadow cost frontier (which does incorporate technical inefficiency)

$$c(y,w^s,t;\beta) = \sum_n w_n^s x_n \cdot \exp\{-\tau\} = \exp\{-\tau\} \cdot c^s(y,w^s,t;\beta). \tag{8.3.8}$$

A lot of tedious algebra leads to the following expression for actual expenditure $E = w^T x$:

$$\begin{aligned} \ln E &= \ln c^s(y,w^s,t;\beta) + \ln G(y,w^s,t;\beta) \\ &= \ln c(y,w^s,t;\beta) + \ln G(y,w^s,t;\beta) + \tau, \end{aligned} \tag{8.3.9}$$

where

$$G(y,w^s,t;\beta) = \sum_n S_n^s \cdot \exp\{\xi_n\},$$

$$\begin{aligned} S_n^s &= \frac{w_n^s x_n}{c^s(y,w^s,t;\beta)} \\ &= \frac{\partial \ln c^s(y,w^s,t;\beta)}{\partial \ln w_n^s} \\ &= \frac{\partial \ln c(y,w^s,t;\beta)}{\partial \ln w_n^s}. \end{aligned}$$

Totally differentiating equation (8.3.9) and the equality $E = w^T x$, and setting the two expressions equal, gives an initial decomposition of dual productivity change

$$\begin{aligned} T\dot{F}P &= [1 - \varepsilon^s(y,w^s,t;\beta)]\dot{y} - \dot{c}^s(y,w^s,t;\beta) \\ &\quad + \sum_n (S_n \dot{w}_n - S_n^s \dot{w}_n^s) - \frac{\partial \ln G(y,w^s,t;\beta)}{\partial t} - \frac{\partial \tau}{\partial t}, \end{aligned} \tag{8.3.10}$$

where $\varepsilon^s(y, w^s, t; \beta) = \partial \ln c^s(y, w^s, t; \beta)/\partial \ln y$. This expression can be rewritten in terms of the cost frontier $c(y, w, t; \beta)$ [which is dual to the original production frontier $f(x, t; \beta)$], on which both technical and allocative efficiency hold, as

$$T\dot{F}P = [1 - \varepsilon(y,w,t;\beta)]\dot{y} - \dot{c}(y,w,t;\beta) - \frac{\partial \tau}{\partial t} + \rho(y,w,t;\beta), \tag{8.3.11}$$

where the cost elasticity $\varepsilon(y, w, t; \beta) = \partial \ln c(y, w, t; \beta)/\partial \ln y$ measures returns to scale on the cost frontier $c(y, w, t; \beta)$, $[-\dot{c}(y, w, t; \beta)]$ measures technical change on the cost frontier $c(y, w, t; \beta)$, $[-\partial\tau/\partial t]$ measures technical efficiency change, and

$$\begin{aligned} \rho(y,w,t;\beta) &= [\dot{c}(y,w,t;\beta) - \dot{c}^s(y,w^s,t;\beta)] \\ &\quad + [\varepsilon(y,w,t;\beta) - \varepsilon^s(y,w^s,t;\beta)]\dot{y} \\ &\quad + \sum_n (S_n \dot{w}_n - S_n^s \dot{w}_n^s) - \frac{d}{dt}\left[\ln\left\{\sum_n S_n^s \left(\frac{w_n}{w_n^s}\right)\right\}\right] \end{aligned}$$

measures the contribution of change in input allocative inefficiency to productivity change. The term $\rho(y, w, t; \beta) = 0$ if the input mix is allocatively efficient, in which case $w_n^s = w_n \Leftrightarrow \xi_n = 0$, $n = 2, \ldots, N$. However $\rho(y, w, t; \beta) = 0$ is also possible in the presence of input allocative inefficiency. The first term in $\rho(y, w, t; \beta)$ is zero if technical change is the same when measured on $c(y, w, t; \beta)$ as when measured on $c^s(y, w^s, t; \beta)$. The second term is zero if scale economies are the same on $c(y, w, t; \beta)$ and $c^s(y, w^s, t; \beta)$. The third term is zero if $\dot{w}_n = \dot{w}_n^s = \dot{w}_k = \dot{w}_k^s\ \forall n, k$. The final term is zero if the (w_n/w_n^s) are independent of time. Each of these cases, and all of them collectively, is consistent with the presence of input allocative inefficiency.

The dual productivity change decomposition given in equation (8.3.11) is similar to the decomposition given in equation (8.3.3). They contain identical scale and technical change components. However the technical efficiency change component in equation (8.3.11) measures the contribution of change in the *magnitude* of technical inefficiency, whereas that in equation (8.3.3) measures the contribution of change in the *cost* of technical inefficiency. Finally, the two remaining components [$\rho(y, w, t; \beta)$ in equation (8.3.11) and $\{\Sigma_n[S_n - S_n(y, w, t; \beta)]\dot{w}_n - \partial u_A/\partial t\}$ in equation (8.3.3)] provide different characterizations of the contribution of change in allocative efficiency. The allocative efficiency component in equation (8.3.11) measures the contribution of changes in the *magnitudes* of allocative inefficiency represented by the parameter vector ξ, whereas the allocative efficiency component in equation (8.3.3) measures the contribution of changes in the *cost* of allocative inefficiency. Finally, if persistent technical and allocative efficiency hold, then the final two components in equation (8.3.11) disappear, and our dual productivity change decomposition collapses to that of Denny, Fuss, and Waverman (1981; 195).

8.3.2 Estimation and Decomposition

It is possible to base estimation and decomposition of dual productivity change on equations (8.3.1)–(8.3.3). All that is required is to follow the estimation strategy outlined in Section 4.3.2, preferably with a flexible specification for the cost frontier and with t added as a regressor to serve as a proxy for technical change. After estimation, estimated values of the technology parameters and the parameters of the technical and allocative inefficiency distributions can be used to implement the productivity change decomposition in equation (8.3.3). However, as we mentioned in Chapter 4, and repeated beneath equation (8.3.5), estimation of the model is complicated by its error structure.

Here we consider estimation and decomposition of dual productivity change based on equations (8.3.6)–(8.3.11). We begin by assigning a translog functional form to the shadow cost frontier $c^s(y, w^s, t; \beta)$, which we write as

$$\begin{aligned}\ln c^s(y_{it}, w_{it}^s, t;\beta) &= \beta_o + \beta_y \ln y_{it} + \sum_n \beta_n \ln w_{nit}^s + \beta_t t \\ &+ \frac{1}{2}\beta_{yy}[\ln y_{it}]^2 + \frac{1}{2}\sum_n\sum_k \beta_{nk}[\ln w_{nit}^s][\ln w_{kit}^s] \\ &+ \frac{1}{2}\beta_{tt}t^2 + \sum_n \beta_{yn}[\ln y_{it}][\ln w_{nit}^s] \\ &+ \beta_{yt}[\ln y_{it}]\mathrm{t} + \sum_n \beta_{nt}[\ln w_{nit}^s]t. \end{aligned} \tag{8.3.12}$$

The corresponding shadow input cost share equations are

$$S_{nit}^s(y_{it}, w_{it}^s, t;\beta) = \beta_n + \sum_k \beta_{nk}\ln w_{kit}^s + \beta_{yn}\ln y_{it} + \beta_{nt}t, \quad n = 1, \ldots, N. \tag{8.3.13}$$

Using equation (8.3.9) and the expressions beneath it, actual expenditure and the actual input cost share equations can be expressed as

$$\begin{aligned}\ln E_{it} &= \ln c^s(y_{it}, w_{it}^s, t;\beta) \\ &+ \ln\left[\sum_n S_{nit}^s(y_{it}, w_{it}^s, t;\beta)\cdot\exp\{\xi_{nit}\}\right] + \tau_{it} + v_{it},\end{aligned} \tag{8.3.14}$$

$$S_{nit} = \frac{S_{nit}^*(y_{it}, w_{it}^s, t;\beta)\cdot\exp\{\xi_{nit}\}}{\Sigma_k S_{kit}^*(y_{it}, w_{it}^s, t;\beta)\cdot\exp\{\xi_{kit}\}} + \eta_{nit}, \qquad n = 1, \ldots, N, \tag{8.3.15}$$

where τ_{it} is the magnitude (and the cost) of input-oriented technical inefficiency, and v_{it} and the η_{nit} are classical random-error terms.

The estimation strategy depends on the nature of the assumptions made on τ_{it} and the ξ_{nit}. If both τ_{it} and the ξ_{nit} are assumed to be random, and distributional assumptions are made on them, it is not possible to derive the likelihood function for the system of N equations consisting of equation (8.3.14) and $N - 1$ of the equations (8.3.15). Therefore, to keep the estimation problem tractable, we assume that the allocative inefficiencies (the ξ_{nit}) are fixed parameters that are constant across producers and through time (and so become ξ_n). Although it is somewhat restrictive, the discussion beneath equation (8.3.11) demonstrates that this assumption does not imply that productivity growth is independent of the ξ_n. We also make the Battese and Coelli (1992) distributional assumption that $\tau_{it} = \tau_i \cdot \exp\{-\gamma(t - T)\}$, where $\tau_i \sim \text{iid } N^+(\mu, \sigma_\tau^2)$, and we also assume that $v_{it} \sim \text{iid } N(0, \sigma_v^2)$, $\eta_{it} \sim N(0, \Sigma_\eta)$, with v distributed independently of τ and η, and τ distributed independently of η. With this specification the likelihood function for the preceding system of equations can easily be derived.

Due to the highly nonlinear nature of equations (8.3.14) and (8.3.15), obtaining maximum likelihood estimators is apt to be computationally difficult. However if the estimators can be obtained, estimates of τ_{it} can be calculated from the residuals of equation (8.3.14) using the JLMS procedure. Under this specification for τ_{it} the contribution of technical efficiency change to productivity change is given by $[-\partial\tau_{it}/\partial t] = \gamma\cdot\tau_{it}$. The remaining components of productivity change identified in equation (8.3.11) can be obtained directly from the estimated parameters in the expenditure equation (8.3.14).

Estimation can be simplified somewhat by using the following two-step procedure. In the first step the input cost share equations are estimated using ITSUR (which does not require distributional assumptions on η). The estimates of S_n^s and ξ_n can be used to obtain an estimate of $[\Sigma_n S^s_{nit}(y_{it}, w^s_{it}, t; \beta)\cdot\exp\{\xi_n\}]$. In the second step maximum likelihood techniques are used to estimate the transformed expenditure equation

$$\ln E_{it} - \ln\left[\sum_n S^s_{nit}(y_{it}, w^s_{it}, t;\beta)\cdot\exp\{\xi_n\}\right] = \ln c^s(y_{it}, w^s_{it}, t;\beta) + \tau_{it} + v_{it}. \quad (8.3.16)$$

Since $\ln c^s(y_{it}, w^s_{it}, t; \beta)$ depends on the ξ_n via the w^s_{nit}, further simplification in estimation can be achieved by treating w^s_{nit} as the price of x_{nit} calculated using the relationship $w^s_{nit} = w_{nit}\cdot\exp\{\xi_n\}$, and not estimating the ξ_n again in the second step. Under this scenario, estimation of the transformed expenditure equation in the second step is not different from estimating technical inefficiency in the Battese and Coelli (1992) model.

8.3.3 An Extension to Multiple Outputs

The analysis, estimation, and decomposition of dual productivity change in Sections 8.3.1 and 8.3.2 is based on the assumption that only a single output is produced. In this section we extend the analysis to the multiple-output case. Our exposition is brief, because the analytical and econometric extensions are straightforward and, more importantly, because our interest centers on the introduction of efficiency change into a decomposition of productivity change, and the extension to multiple outputs is of secondary interest.

The key difference between the decomposition of productivity change in the multiple-output case and in the single-output case lies in the definition of productivity change. When there are M outputs $y = (y_1, \ldots, y_M) \geq 0$, the Divisia index of productivity change given in equation (8.2.4) becomes

$$\begin{aligned} T\dot{F}P &= \dot{Y} - \dot{X} \\ &= \sum_m R_m \dot{y}_m - \sum_n S_n \dot{x}_n, \end{aligned} \quad (8.3.17)$$

where $R_m = p_m y_m / R$ is the observed revenue share of output y_m, p_m is the price of output y_m, and $R = \Sigma_m p_m y_m$ is total revenue. If the scalar output y in the deterministic cost frontier given in equation (8.3.1) is replaced by the output vector $y = (y_1, \ldots, y_M)$, then the analysis of Section 8.3.1 can be adapted to the multiple-output case.

Adapting the error components analysis in equations (8.3.1)–(8.3.3) to the multiple-output case leads to the following expression for productivity change:

$$\begin{aligned} T\dot{F}P = {} & [1 - \varepsilon(y, w, t; \beta)]\dot{Y}^c - \dot{c}(y, w, t; \beta) + \sum_n [S_n - S_n(y, w, t; \beta)]\dot{w}_n \\ & + (\dot{Y} - \dot{Y}^c) - \frac{\partial u_T}{\partial t} - \frac{\partial u_A}{\partial t}, \end{aligned} \quad (8.3.18)$$

where

$$\dot{Y}^c = \sum_m \left[\frac{\varepsilon_m(y, w, t; \beta)}{\varepsilon(y, w, t; \beta)}\right]\dot{y}_m$$

provides a measure of aggregate output growth using efficient cost elasticity weights in place of revenue share weights, the $\varepsilon_m(y, w, t; \beta) = \partial \ln c(y, w, t; \beta)/\partial \ln y_m$ are elasticities of minimum cost with respect to y_m, $m = 1, \ldots, M$, and $\varepsilon(y, w, t; \beta) = \Sigma_m \varepsilon_m(y, w, t; \beta)$ is the dual multiproduct measure of scale economies.

Expression (8.3.18) appears in Bauer (1990a; 292) and, in simplified form, in Denny, Fuss, and Waverman (1981; 197). It decomposes productivity change into a scale effect, a technical change effect, an input allocative efficiency effect, and a pair of efficiency cost effects, as in the single-output version given in equation (8.3.3). The effect $(\dot{Y} - \dot{Y}^c)$ is new, and captures the impact on productivity change of departures from marginal cost pricing, or of departures from equiproportionate markups over marginal cost pricing, since $(\dot{Y} - \dot{Y}^c) = 0$ if $p_m = \theta\, \partial c(y, w, t; \beta)/\partial y_m$, $m = 1, \ldots, M$. This term does not appear in equation (8.3.3), apparently because there is no question of appropriate output weights in the single-output case, despite the fact that $p \neq \partial c(y, w, t; \beta)/\partial y$ remains possible in the single-output case.

In principle productivity change can be estimated and decomposed in the multiple-output composed error framework; Bauer (1990a) did so under several restrictive assumptions. However, as we indicated previously, the complicated error structure makes the model exceedingly difficult to estimate, regardless of the number of outputs, and so we now consider a multiproduct model in which technical and input allocative inefficiency are modeled parametrically rather than through error components. This approach extends the analysis and estimation in equations (8.3.6)–(8.3.16) to the multiple-output case.

Equation (8.3.11) provides a decomposition of dual productivity change in the single-output case. The extension to multiple outputs is given by

$$T\dot{F}P = [1-\varepsilon(y,w,t;\beta)]\dot{Y}^c - \dot{c}(y,w,t;\beta) - \frac{\partial \tau}{\partial t} + \rho(y,w,t;\beta), \qquad (8.3.19)$$

where

$$\rho(y,w,t;\beta) = [\dot{c}(y,w,t;\beta) - \dot{c}^s(y,w^s,t;\beta)] + \{[\varepsilon(y,w,t;\beta)]\dot{Y}^c - [\varepsilon^s(y,w^s,t;\beta)]\dot{Y}^{cs}\} + (\dot{Y} - \dot{Y}^c) + \sum_n [S_n\dot{w}_n - S_n^s\dot{w}_n^s] - \frac{d}{dt}\left[\ln\left\{\sum_n S_n^s\left(\frac{w_n}{w_n^s}\right)\right\}\right]$$

and $\dot{Y}^{cs} = \Sigma_m[\varepsilon_m^s(y,w^s,t;\beta)/\varepsilon^s(y,w^s,t;\beta)]\dot{y}_m$ provides a measure of aggregate output growth using shadow input prices to construct cost elasticity weights. All other terms are as defined before.

Equation (8.3.19) decomposes dual productivity change into a scale term, a technical change term, a technical efficiency change term, and an input allocative efficiency change term, exactly as equation (8.3.11) does in the single-output case. The technical change and technical efficiency change terms are unchanged from their single-output versions. The scale term is a straightforward generalization of the single-output scale term. The input allocative efficiency change term $\rho(y,w,t;\beta)$ is a multiple-output generalization of the single-output expression introduced beneath equation (8.3.11), with one exception. The component $(\dot{Y} - \dot{Y}^c)$ is new, and captures disproportionate departures from marginal cost pricing. Thus, as in the single-output case, the effect of input allocative inefficiency on productivity change has several components: (i) a technical change component, which is zero if technical change is the same on $c(y,w,t;\beta)$ and $c^s(y,w^s,t;\beta)$; (ii) a scale component, which is zero if returns to scale are the same on $c(y,w,t;\beta)$ and $c^s(y,w^s,t;\beta)$; (iii) a disproportionate departures from marginal cost pricing term; and (iv) a miscellaneous term, which is zero if actual and shadow input cost shares are the same and the (w_n/w_n^s) are constant for all $n = 1, \dots, N$.

Estimation and decomposition in the multiple-output case proceed exactly as in the single-output case discussed in Section 8.3.2. All that is required is to replace the scalar output y_{it} with a vector of outputs y_{mit}, $m = 1, \dots, M$, in the translog cost frontier system given in equations (8.3.12)–(8.3.16).

8.4 A DUAL (PROFIT FRONTIER) APPROACH

A dual profit-maximizing framework is not widely used to analyze, estimate, and decompose productivity change, although it is possible to do so. Two approaches are available. It is possible to use a primal framework and invoke the first-order conditions for profit maximization to derive an expression for productivity change. Alternatively, it is possible to use a dual framework to begin with. In this section we follow the latter approach. The analytical framework we develop in Section 8.4.1 is structurally similar to the analytical framework we developed in Section 8.3.1 for use in a cost-minimizing environment. The estimation and decomposition procedures we develop in Section 8.4.2 are also similar to the empirical procedures we developed in Section 8.3.2 for use in a cost-minimizing environment. Finally, the extension to multiple outputs in Section 8.4.3 is also similar to the extension we developed in Section 8.3.3.

8.4.1 The Analytical Framework

A conventional profit frontier $\pi(p,w,t;\beta)$ is defined as the solution to the constrained maximization problem

$$\pi(p,w,t;\beta) = \max_{y,x}\{py - w^Tx:\ y = f(x,t;\beta)\}, \qquad (8.4.1)$$

where $f(x,t;\beta)$ is the deterministic kernel of a stochastic production frontier, $y \geq 0$ is scalar output with price $p > 0$, $x = (x_1, \dots, x_N) \geq 0$ is an input quantity vector with price vector $w = (w_1, \dots, w_N) > 0$, and β is a vector of technology parameters to be estimated. Technical inefficiency can be introduced in either input-oriented or output-oriented form; in output-oriented form it is introduced by scaling output y to technically efficient output $y\cdot\exp\{u\}$ with $u \geq 0$. Input allocative inefficiency can be introduced by replacing input prices with shadow input prices $w^s = [w_1, w_2\cdot\exp\{-\xi_2\}, \dots, w_N\cdot\exp\{-\xi_N\}]$. The corresponding shadow profit frontier is

$$\pi^s(p^s,w^s,t;\beta) = \max_{y,x}\{p^s y\cdot\exp\{u\} - w^{sT}x:\ y\cdot\exp\{u\} = f(x,t;\beta)\}, \qquad (8.4.2)$$

where $p^s = p\cdot\exp\{-u\}$ is the shadow price of technically efficient output $y\cdot\exp\{u\}$.

From the expression for $\pi^s(p^s, w^s, t; \beta)$ it follows that the rate of change of shadow profit can be written as

$$\dot{\pi}^s = \left(\frac{\partial \ln \pi^s}{\partial \ln p^s}\right)\dot{p}^s + \sum_n \left(\frac{\partial \ln \pi^s}{\partial \ln w_n^s}\right)\dot{w}_n^s + \frac{\partial \ln \pi^s}{\partial t}$$
$$= \frac{1}{\pi^s}\left\{py\dot{p}^s - c^s \sum_n S_n^s \dot{w}_n^s\right\} + \frac{\partial \ln \pi^s}{\partial t}, \tag{8.4.3}$$

since $\partial\pi^s/\partial p = y$ and $\partial\pi^s/\partial w_n^s = -x_n$, $n = 1, \ldots, N$, by Hotelling's lemma. Since shadow profit can also be expressed as $\pi^s = py - \Sigma_n w_n^s x_n$, we have

$$\dot{\pi}^s = \frac{py}{\pi^s}\{\dot{y} + \dot{p}\} - \frac{c^s}{\pi^s}\sum_n \{S_n^s \dot{x}_n + S_n^s \dot{w}_n^s\}. \tag{8.4.4}$$

Setting equations (8.4.3) and (8.4.4) equal yields, after some algebraic manipulations,

$$\pi^s\left(\frac{\partial \ln \pi^s}{\partial t}\right) = py\left(\dot{y} - \sum_n S_n \dot{x}_n\right) + py\sum_n S_n \dot{x}_n$$
$$-c^s \sum_n S_n^s \dot{x}_n + py\left(\frac{\partial u}{\partial t}\right)$$
$$= py(T\dot{F}P) - py(\varepsilon^s - 1)\sum_n S_n^s \dot{x}_n$$
$$-py\sum_n (S_n^s - S_n)\dot{x}_n + py\left(\frac{\partial u}{\partial t}\right), \tag{8.4.5}$$

where $\varepsilon^s = \Sigma_n(\partial \ln y/\partial \ln x_n)$ evaluated using the first-order conditions for profit maximization, so that $\varepsilon^s = \Sigma_n w_n^s x_n/py = c^s/py$. Alternatively, $\varepsilon^s = [1 - (\partial \ln \pi^s/\partial \ln p^s)^{-1}]$. Solving equation (8.4.5) for the measure of productivity change yields

$$T\dot{F}P = \frac{\pi^s}{py}\frac{\partial \ln \pi^s}{\partial t} + (\varepsilon^s - 1)\sum_n S_n^s \dot{x}_n + \sum_n (S_n^s - S_n)\dot{x}_n - \frac{\partial u}{\partial t}. \tag{8.4.6}$$

Equation (8.4.6) provides a decomposition of dual productivity change into four components. Applying the envelope theorem to equation (8.4.2) yields $\partial\pi^s/\partial t = p^s \cdot [\partial f(x, t; \beta)/\partial t]$, which in turn implies that $(\pi^s/py)(\partial \ln \pi^s/\partial t) = [f(x, t; \beta)]^{-1} \cdot [\partial f(x, t; \beta)/\partial t]$. Thus the first component on the right-hand side of equation (8.4.6) is equivalent to $T\Delta$ in equations (8.2.5) and (8.2.6). Similarly ε^s and the S_n^s are the same as ε and the S_n in equations (8.2.5) and (8.2.6) when the first-order conditions for profit maximization are imposed. Consequently the second component on the right-hand side of equation (8.4.6) provides a dual measure of the contribution of scale economies to productivity change. Also when $\xi_n = 0\ \forall n$, $S_n^s = S_n$ and the third component on the right-hand side of equation (8.4.6) disappears. Finally the fourth component on the right-hand side of equation (8.4.6) is a dual measure of the contribution of output-oriented technical efficiency change to productivity change, and corresponds to $TE\Delta$ in equations (8.2.5) and (8.2.6).

There is, however, one important difference between the decomposition in equation (8.2.5) and that in equation (8.4.6). In deriving equation (8.2.5) we used quantity data only, and we did not impose any optimization conditions explicitly, whereas in equation (8.4.6) we used both quantity and price data, and we imposed a profit maximization condition in the allocation of the input and output quantities. It is these optimization conditions that enable us to show that the primal and dual productivity change decompositions are identical. In the absence of these conditions, the two decompositions can differ, as they did in the cost-oriented analysis in Section 8.3.

Each of the terms appearing on the right-hand side of equation (8.4.6) can be expressed in terms of the shadow profit function and its derivatives, and so estimation of the shadow profit function enables one to estimate and decompose dual productivity growth. The requisite terms and their expressions are as follows:

$$py = \pi^s\left(\frac{\partial \ln \pi^s}{\partial \ln p^s}\right),$$
$$(\varepsilon^s - 1) = -\left(\frac{\partial \ln \pi^s}{\partial \ln p^s}\right)^{-1},$$
$$S_n^s = \frac{\partial \ln \pi^s/\partial \ln w_n^s}{\Sigma_k(\partial \ln \pi^s/\partial \ln w_k^s)}, \qquad n = 1, \ldots, N,$$
$$S_n = \frac{(\partial \ln \pi^s/\partial \ln w_n^s)(w_n/w_n^s)}{\Sigma_k[\partial \ln \pi^s/\partial \ln w_k^s)(w_k/w_k^s)]}, \qquad n = 1, \ldots, N,$$
$$\dot{x}_n = \frac{\partial^2 \ln \pi^s}{\partial \ln w_n^s \partial t}\left[\frac{\partial \ln \pi^s}{\partial \ln w_n^s}\right]^{-1} + \frac{\partial \ln \pi^s}{\partial t} - \frac{\partial \ln w_n^s}{\partial t}, \qquad n = 1, \ldots, N.$$

The productivity growth decomposition given in equation (8.4.6) can also be expressed in terms of the profit frontier $\pi(p, w, t; \beta)$. In this decomposition we define returns to scale and technical change on the profit frontier and rewrite equation (8.4.6) as

$$T\dot{F}P = \frac{\pi}{py}\frac{\partial \ln \pi}{\partial t} + (\varepsilon - 1)\sum_n S_n \dot{x}_n - \frac{\partial u}{\partial t} + \rho(p, w, t, \xi, u), \tag{8.4.7}$$

where

$$\rho(p,w,t,\xi,u) = \frac{\pi^s}{py}\frac{\partial \ln \pi^s}{\partial t} - \left[\frac{\partial \ln \pi}{\partial \ln p}\right]^{-1}\frac{\partial \ln \pi}{\partial t} + (\varepsilon^s - 1)\sum_n S_n^s \dot{x}_n$$
$$-(\varepsilon - 1)\sum_n S_n \dot{x}_n + \sum_n (S_n^s - S_n)\dot{x}_n.$$

In this formulation dual productivity change is decomposed into a technical change component, a scale component, and an efficiency change component, each defined relative to the profit frontier $\pi(p, w, t; \beta)$, and a miscellaneous term. The miscellaneous term is zero if technical change is the same along $\pi(p, w, t; \beta)$ and $\pi^s(p^s, w^s, t; \beta)$, if returns to scale are the same along $\pi(p, w, t; \beta)$ and $\pi^s(p^s, w^s, t; \beta)$, and if shadow input cost shares coincide with actual input cost shares. Expressions for ε and the S_n can be obtained from the expressions for ε^s and the S_n^s given previously, by substituting π for π^s, w_n for w_n^s, and p for p^s. Note that ρ depends on output-oriented technical inefficiency u, and so $\rho \neq 0$ even in the absence of input allocative inefficiency.

8.4.2 Estimation and Decomposition

In this section we consider estimation of the productivity change components that are derived from the shadow profit function $\pi^s(p^s, w^s, t; \beta)$. A single-output and multiple-input production technology is assumed. Our starting point is the relationship between actual profit $\pi = py - w^T x$ and shadow profit $\pi^s(p^s, w^s, t; \beta)$ in the presence of both technical and allocative inefficiencies. From the relationships developed in Chapters 5 and 6 we have

$$\ln \pi = \ln \pi^s + \ln\left[SR^s - \sum_n SC_n^s \cdot \exp\{\xi_n\}\right], \tag{8.4.8}$$

where

$$\ln \pi^s = \beta_o + \beta_p \ln p^s + \sum_n \beta_n \ln w_n^s + \beta_t t + \frac{1}{2}\beta_{pp}(\ln p^s)^2$$
$$+ \frac{1}{2}\sum_n \sum_k \beta_{nk} \ln w_n^s \ln w_k^s + \frac{1}{2}\beta_{tt} t^2 + \sum_n \beta_{pn} \ln p^s \ln w_n^s$$
$$+ \beta_{pt} \ln p^s t + \sum_n \beta_{nt} \ln w_n^s t, \tag{8.4.9}$$

$$SR^s = \beta_p + \beta_{pp} \ln p^s + \sum_n \beta_{pn} \ln w_n^s + \beta_{pt} t, \tag{8.4.10}$$

$$-SC_n^s = \beta_n + \sum_k \beta_{nk} \ln w_k^s + \beta_{pn} \ln p^s + \beta_{nt} t, \qquad n = 1, \ldots, N, \tag{8.4.11}$$

and $SR^s = p^s y/\pi^s = \partial \ln \pi^s/\partial \ln p^s$ and $SC_n^s = w_n^s x_n/\pi^s = -\partial \ln \pi^s/\partial \ln w_n^s$, $n = 1, \ldots, N$, are shadow profit shares. The shadow profit function $\pi^s(p^s, w^s, t; \beta)$ is homogeneous of degree +1 in (p^s, w^s), which together with symmetry imposes the parameter restrictions $\beta_p + \Sigma_n \beta_n = 1$, $\beta_{pp} + \Sigma_n \beta_{pn} = 0$, $\beta_{pn} + \Sigma_k \beta_{nk} = 0\ \forall n$, $\beta_{pt} + \Sigma_n \beta_{nt} = 0$, $\beta_{nk} = \beta_{kn}\ \forall n \neq k$. The actual profit share equations are

$$SR = \frac{py}{\pi} = \frac{SR^s}{SR^s - \Sigma_k SC_k^s \cdot \exp\{\xi_k\}}, \tag{8.4.12}$$

$$SC_n = \frac{w_n x_n}{\pi} = \frac{SC_n^s \cdot \exp\{\xi_n\}}{SR^s - \Sigma_k SC_k^s \cdot \exp\{\xi_k\}}, \qquad n = 1, \ldots, N. \tag{8.4.13}$$

The problem is to estimate the system of equations (8.4.12) and (8.4.13). However even if we assume that the allocative inefficiencies are fixed parameters invariant through time and across producers, this system cannot be estimated in the same way as the corresponding cost system can be estimated. The reason for this is that the profit share equations are affected by the presence of technical inefficiency via the output shadow price p^s. That is, SR^s and the SC_n^s depend on u, and since u is random it is virtually impossible to derive the distribution of the error vector (or simplify the variance–covariance matrix) in the system of equations (8.4.12) and (8.4.13).

Here we suggest three methods of estimation, although none of these methods is entirely satisfactory. The first approach uses a very restrictive specification of production technology, which simplifies matters considerably at an equally considerable sacrifice of flexibility. The second and third approaches retain flexibility, but they are difficult to estimate.

We begin by assuming a Cobb–Douglas functional form, in which case the deterministic production frontier is

$$\ln y = b_o + \sum_n b_n \ln x_n + b_t t - u, \tag{8.4.14}$$

and the corresponding shadow profit frontier is

$$\ln \pi^s = \beta_o + \beta_p \ln p^s + \sum_n \beta_n \ln w_n^s + \beta_t t, \tag{8.4.15}$$

where $\beta_p = 1 - \Sigma_n \beta_n$, $\beta_n = -b_n/(1 - \Sigma_n b_n)$, $n = 1, \ldots, N$, and $\Sigma_n b_n < 1$. In this simplified model the input profit share equations become

$$SC_n = \frac{\beta_n \cdot \exp\{\xi_n\}}{(1 - \Sigma_k \beta_k) + \Sigma_k \beta_k \cdot \exp\{\xi_k\}}, \qquad n = 1, \ldots, N, \tag{8.4.16}$$

and the output profit share equation is $SR = 1 - \Sigma_n SC_n$.

It is clear from equations (8.4.16) that even though we have restricted ourselves to the Cobb–Douglas case, the allocative errors ξ_n do not appear linearly in the actual profit share equations. What this means is that if one considers the ξ_n as random variables, estimation of the actual profit share equations in (8.4.16) will still be difficult. This problem can, however, be avoided if one uses the production frontier and the first-order conditions for profit maximization, and considers the system of equations

$$\begin{aligned} &\ln y - b_o - \sum_n b_n \ln x_n - b_t t = v - u, \\ &\ln x_n - \sum_k b_k \ln x_k - b_t t - \ln p + \ln w_n - \ln b_n - \ln b_o = \xi_n - u, \\ &\qquad n = 1, \ldots, N. \end{aligned} \tag{8.4.17}$$

In this specification $v \sim$ iid $N(0, \sigma_v^2)$ is a random-noise error component and $u \geq 0$ represents output-oriented technical inefficiency. If distributional assumptions are imposed on u and the ξ_n, this system can be estimated following procedures outlined in Kumbhakar (1987, 1990). Once the parameters are estimated [including u from the generalization of the JLMS procedure given in Kumbhakar (1987)], all of the productivity change components given at the end of Section 8.4.1 can be derived.

Next we return to the original system of profit share equations given in equations (8.4.12) and (8.4.13), which are based on a flexible translog specification. It is important to note that even if u (embedded in p^s) and the ξ_n are time invariant, productivity change is nonetheless not independent of technical or allocative inefficiencies. This is because π^s, ε^s, and the S_n^s all depend on u and the ξ_n. However if the allocative inefficiencies ξ_n are modeled as fixed parameters, invariant across producers and through time, and if the technical inefficiencies u are modeled as time-invariant parameters, then it is possible to estimate the system (8.4.12) and (8.4.13) using a nonlinear ITSUR approach. However if the number of producers is large, estimation is still likely to be a problem.

The final approach to estimation we consider is based on an approach we introduced in Chapter 6. We begin by rewriting the shadow profit share equations (8.4.10) and (8.4.11) as $SR^s = SR^o - u\beta_{pp}$ and $-SC_n^s = -SC_n^o - u\beta_{pn}$, $n = 1, \ldots, N$, where

$$\begin{aligned} SR^o &= \beta_p + \beta_{pp} \ln p + \sum_n \beta_{pn} \ln w_n^s + \beta_{pt} t, \\ -SC_n^o &= \beta_n + \sum_k \beta_{nk} \ln w_k^s + \beta_{pn} \ln p + \beta_{nt} t, \qquad n = 1, \ldots, N. \end{aligned}$$

Then the denominators in equations (8.4.12) and (8.4.13) become

$$\begin{aligned} SR^s - \sum_k SC_k^s \cdot \exp\{\xi_k\} &= SR^o - \sum_k SC_k^o \cdot \exp\{\xi_k\} - uA_o \\ &= D_o - uA_o, \end{aligned}$$

where $A_o = \beta_{pp} + \Sigma_k \beta_{pk} \cdot \exp\{\xi_k\}$. Using these results, the actual profit share equations can be rewritten as

$$\begin{aligned} &\frac{D_o SR - SR^o}{A_o SR - \beta_{pp}} = u, \\ &\frac{D_o SC_n - SC_n^o}{A_o SC_n + \beta_{pn}} = u, \qquad n = 1, \ldots, N. \end{aligned} \tag{8.4.18}$$

Next we eliminate u from these actual profit share equations by subtracting each input profit share equation from the output profit share equation to obtain

$$\frac{D_o SR - SR^o}{A_o SR - \beta_{pp}} - \frac{D_o SC_n - SC_n^o}{A_o SC_n + \beta_{pn}} = \eta_n, \qquad n = 1, \ldots, N, \tag{8.4.19}$$

where the η_n are classical error terms. There is no loss of information in using N of $N + 1$ profit share equations, since only N share equations are independent. Assuming that the ξ_n are producer- and time-invariant fixed parameters, and that the error vector η has zero mean and constant variance–covariance matrix, the system of equations (8.4.19) can be estimated using the nonlinear ITSUR procedure. As can be seen from the expressions for SR^o and the SC_n^o, the system in (8.4.19) is highly nonlinear. The only comfort is that the parameters are estimated without making any distributional assumptions. Technical inefficiency u is not estimated from the system in (8.4.19). It can, however, be estimated from (8.4.18), from the mean of the left-hand sides of the $N + 1$ profit share equations using the estimated parameters from equations (8.4.19). Since some of the parameters in π^s do not appear in the profit share equations, and are therefore not estimated, we need another step to estimate the remaining parameters. This step involves use of the shadow profit function in (8.4.9), after substituting the parameters already estimated, to obtain a regression equation of the form $Z = \beta_o + \beta_t t + (1/2)\beta_{tt}t^2$, where $Z = \ln \pi^s - \ln[SR^s - \Sigma_n SC_n^s \cdot \exp\{\xi_n\}]$ – all terms on the right-hand side of (8.4.9) except the intercept and the terms involving t and t^2.

8.4.3 An Extension to Multiple Outputs

As in Section 8.3.3 we define productivity change in the multiple-output case as $T\dot{F}P = \dot{Y} - \dot{X} = \Sigma_m R_m \dot{y}_m - \Sigma_n S_n \dot{x}_n$. Again we work with a profit function $\pi(p, w, t; \beta)$ and a shadow profit function $\pi^s(p^s, w^s, t; \beta)$, where now $p = (p_1, \ldots, p_M)$ and $p^s = p \cdot \exp\{-u\}$. Totally differentiating the shadow profit function yields

$$\begin{aligned}\dot{\pi}^s &= \sum_m \left(\frac{\partial \ln \pi^s}{\partial \ln p_m^s}\right)\dot{p}_m^s + \sum_n \left(\frac{\partial \ln \pi^s}{\partial \ln w_n^s}\right)\dot{w}_n^s + \frac{\partial \ln \pi^s}{\partial t} \\ &= \frac{1}{\pi^s}\left\{R\sum_m R_m \dot{p}_m^s - C^s \sum_n S_n^s \dot{w}_n^s\right\} + \frac{\partial \ln \pi^s}{\partial t},\end{aligned} \tag{8.4.20}$$

where $R = \Sigma_m p_m y_m$ and $R_m = p_m y_m / R$, $m = 1, \ldots, M$. Since shadow profit can also be expressed as $\pi^s = \Sigma_m p_m y_m - \Sigma_n w_n^s x_n$, we obtain

$$\dot{\pi}^s = \frac{R}{\pi^s}\left\{\sum_m R_m \dot{y}_m + \sum_m R_m \dot{p}_m\right\} - \frac{C^s}{\pi^s}\sum_n \{S_n \dot{x}_n + S_n \dot{w}_n^s\}. \tag{8.4.21}$$

Equating equations (8.4.20) and (8.4.21) yields

$$\begin{aligned}\pi^s\left(\frac{\partial \ln \pi^s}{\partial t}\right) &= R\left(\sum_m R_m \dot{y}_m - \sum_n S_n \dot{x}_n\right) \\ &\quad + R\sum_n S_n \dot{x}_n - C^s \sum_n S_n \dot{x}_n + \frac{R\partial u}{\partial t} \\ &= R(T\dot{F}P) - R(\varepsilon^s - 1)\sum_n S_n^s \dot{x}_n \\ &\quad - R\sum_n \{S_n^s - S_n\}\dot{x}_n + \frac{R\partial u}{\partial t},\end{aligned} \tag{8.4.22}$$

from which we obtain the following expression for dual productivity change:

$$T\dot{F}P = \frac{\pi^s}{R}\frac{\partial \ln \pi^s}{\partial t} + (\varepsilon^s - 1)\sum_n S_n^s \dot{x}_n + \sum_n \{S_n^s - S_n\}\dot{x}_n - \frac{\partial u}{\partial t}, \tag{8.4.23}$$

which is identical to the single-output productivity change decomposition given in equation (8.4.6), apart from the replacement of py with $R = \Sigma_m p_m y_m$. The interpretation of this decomposition, and of its relationship to the primal decomposition in equations (8.2.5) and (8.2.6), is essentially unchanged.

One feature of equation (8.4.23) warrants mention, because it has no counterpart in the single-output case. The first component on the right-hand side of equation (8.4.23) can be rewritten as

$$\begin{aligned}\frac{\pi^s}{R}\frac{\partial \ln \pi^s}{\partial t} &= \frac{1}{R}\frac{\partial \pi^s}{\partial t} \\ &= \frac{1}{R}\sum_m p_m\left(\frac{\partial y_m}{\partial t}\right) \\ &= \sum_m R_m\left(\frac{\partial \ln y_m}{\partial t}\right).\end{aligned} \tag{8.4.24}$$

Thus the technical change component of productivity change in equation (8.4.23) can also be interpreted as a revenue share weighted average of output-specific rates of technical change. Equation (8.4.24) is thus a dual multiple-output version of the primal technical change component $T\Delta$ appearing in equation (8.2.2).

Estimation is difficult in the single-output case, as we mentioned in Section 8.4.2. The problems there involved nonlinearities rather than the number of inputs and outputs, so adding additional outputs to the model causes no serious additional complications. Nonetheless, it does not make estimation any easier.

8.5 A GUIDE TO THE LITERATURE

A good introduction to index number approaches to productivity change measurement is provided by Balk (1998). More detailed surveys are provided by Diewert (1981a, b, 1987, 1992); these four papers, and others as well, are collected in Diewert and Nakamura (1993).

The primal approach to the econometric estimation of productivity change originated with Solow (1957), who assumed constant returns to scale and technical efficiency, and associated productivity change with technical change. The fact that Solow found such a large residual [which he called productivity growth and Abramovitz (1956) called "a measure of our ignorance"] in the United States during the first half of the twentieth century led to the development of growth accounting. The objective of growth accounting was to eliminate various errors of aggregation and measurement, in an effort to minimize the unexplained residual and to arrive at a more accurate measure of productivity change. Influential work of Denison and Jorgenson and Griliches is contained in a special issue of the *Survey of Current Business* (1972).

The next development in the econometric estimation of productivity change resulted from the application of duality theory to the problem. Ohta (1974) derived the relationships between primal and dual cost measures of scale economies and technical change. Caves, Christensen, and Swanson (1980), Denny, Fuss, and Waverman (1981), and Nadiri and Schankerman (1981) used flexible cost functions to estimate and decompose dual productivity change. These last two papers, and several other productivity studies exploiting duality theory, appear in an excellent conference volume edited by Cowing and Stevenson (1981). A closely related body of literature concerns the measurement, estimation, and decomposition of productivity change in the presence of what Berndt and Fuss (1986) call "temporary equilibrium," due perhaps to the presence of quasi-fixed inputs. Early contributions to this literature include Caves, Christensen, and Swanson (1981), Hulten (1986), Morrison (1985, 1986), and Schankerman and Nadiri (1986), and a comprehensive survey is provided by Morrison (1993).

The introduction of efficiency change as a source of productivity change was pioneered by Nishimizu and Page (1982), who used a deterministic translog production frontier to decompose productivity change in Yugoslavia into technical change and technical efficiency change. Much later Bauer (1990a) merged the work cited in the previous paragraph with that of Nishimizu and Page to derive detailed primal and dual (cost) decompositions of productivity change. Much of the material in Sections 8.2 and 8.3 is based on Bauer's work. Despite the fact that Bauer's work is nearly a decade old, little subsequent effort has been devoted to the introduction of efficiency change into econometric models of productivity change. Surveys of the literature, both econometric and noneconometric, are provided by Grosskopf (1993) and Lovell (1996). More recently interest in an econometric approach to the construction of Malmquist-type productivity indexes has grown rapidly. Noteworthy examples include Atkinson and Cornwell (1998b), who contrast frontier and nonfrontier approaches; Atkinson and Primont (1998), who use shadow cost and distance functions; Coelli (1997), who uses a stochastic production frontier approach; and Fuentes, Grifell-Tatjé, and Perelman (1997), who use stochastic distance functions.