Technologies, Markets and Behaviour: Some Implications for Estimating Efficiency and Productivity Change

C. J. O'Donnell

Centre for Efficiency and Productivity Analysis, School of Economics,
The University of Queensland, Brisbane, 4072, Australia

Paper presented at 58th Annual Conference of the Australian Agricultural and Resource Economics Society, Port Macquarie, 4-7 February, 2014

Abstract

Most productivity indexes can be exhaustively decomposed into measures of technical change and efficiency change. Estimating these components usually involves the use of data envelopment analysis (DEA) or stochastic frontier analysis (SFA) models. This paper shows how assumptions concerning technologies, markets and firm behaviour can be used to frame these models. The paper explains that the assumptions underpinning common DEA models are rarely, if ever, true. On the other hand, the assumptions underpinning basic SFA models are almost always true. The parameters of basic SFA models can be estimated using ordinary least squares and two-stage least squares methods. More complex SFA models can be estimated using maximum likelihood methods. Unfortunately, the assumptions underpinning some of these more complex models are generally not true. This has important implications for estimating the drivers of productivity change. To illustrate, the paper uses common least squares and maximum likelihood methods to estimate the drivers of productivity change in U.S. agriculture. As expected, the different estimators lead to qualitatively different estimates of the efficiency change components productivity change.

---

1. Introduction

Measures of productivity can be calculated without knowing anything about technologies, markets or behaviour. However, if these measures are to inform economic decision-making, then, among other things, decision-makers need to know the maximum productivity that is technically feasible, and the reasons why some firms[1] might rationally choose to be less productive than others. Estimating the maximum productivity that is possible requires knowledge of technologies. Explaining the choices made by firms involves knowledge of markets and firm optimising behaviour. This paper explains how different assumptions concerning technologies, markets and behaviour can be used to guide the estimation of efficiency and productivity change.

The starting point is to define exactly what is meant by the term *technology*. In this paper, a technology is defined as a technique, method or process for transforming inputs into outputs (e.g., a technique for manipulating an organism's genome, or a process for turning grapes into wine). For all practical intents and purposes, it is convenient to think of a technology as a book of instructions. It is common to make assumptions about technologies by way of assumptions about what they can and cannot produce. Common assumptions are:

A1: there is a limit to what can be produced using a finite amount of inputs (output sets are bounded);

A2: a strictly positive amount of at least one input is needed in order to produce a strictly positive amount of any output (inputs are weakly essential);

A3: if an input vector can be used to produce a particular output vector then it can also be used to produce a scalar contraction of that output vector (outputs are weakly disposable);

A4: if an output vector can be produced using a particular input vector then it can also be produced using a scalar magnification of that input vector (inputs are weakly disposable);

[1] In this paper, the term *firm* is used generically to refer to any decision-making unit, or the manager of that decision making unit. Thus, a firm could be anything from an individual to an economy, and the actions of firms are the actions of firm managers.

A5: the set of outputs that can be produced using a given input vector contains all the points on its boundary (output sets are closed); and

A6: the set of inputs that can produce a given output vector contains all the points on its boundary (input sets are closed).

These weak assumptions are maintained throughout the paper and will not be subjected to empirical testing. The main reason for maintaining these assumptions is that, if they are true, then production possibilities sets can be represented using output and input distance metafunctions. Stronger versions of A3 and A4 that have important implications for these functions are:

A7: it is possible to use the same inputs to produce fewer outputs (outputs are strongly disposable); and

A8: it is possible to produce the same outputs using more inputs (inputs are strongly disposable).

If A1–A8 are true, then the output and input distance metafunctions are nonnegative (NN), nondecreasing (ND) and homogeneous of degree one (HD1) in outputs and inputs respectively. This means they can be used to construct proper[2] output quantity and input quantity indexes. This paper shows how they can also be used to construct a new proper total factor productivity (TFP) index. Special cases of this new index include the Färe-Primont TFP index defined by O'Donnell (2012*c*) and the geometric Young (GY) TFP index defined by O'Donnell (2011). Any one of these indexes can be used to make reliable comparisons involving multiple firms and/or time periods (i.e., they are proper multilateral/multitemporal indexes). If there are only two observations in the dataset, then the GY TFP index collapses to the well-known Törnqvist index.

O'Donnell (2012*a*,*d*) shows that all proper TFP indexes can be decomposed into a measure of technical change and various measures of efficiency change. The main aim of this paper is to find weak assumptions that will enable researchers to obtain consistent

[2]In this paper, an output or input quantity index is said to be proper if and only if it satisfies the identity, transitivity, proportionality, homogeneity, time-space reversal, circularity and weak monotonicity axioms listed in O'Donnell (2012*b*). A TFP index is said to be proper if and only if it can be written as a proper output quantity index divided by a proper input quantity index.

estimates of these components. In practice, estimating technical change and efficiency change usually involves estimating relevant characteristics of the output distance metafunction. If it exists, the output distance metafunction can be written in the form of the stochastic frontier analysis (SFA) model of Aigner et al. (1977) and Meeusen and van den Broeck (1977). The composite errors in this well-known model account for the fact that, *inter alia*, the functional form of the metafunction is rarely known, the variables involved in the production process are usually measured with error, and firms may not be technically efficient. This paper shows that consistent estimation of the components of productivity change hinges on two assumptions:

E1: the composite errors are random variables with the same mean; and

E2: the composite errors are uncorrelated with the explanatory variables.

If A1–A6, E1 and E2 are true, then the ordinary least squares (OLS) estimator is a consistent estimator of the slope parameters in the SFA model. Contrary to what is often stated in the literature, no explicit assumptions concerning the functional form of the unknown metafunction are required. However, if the deterministic component of the SFA model is a poor approximation to the unknown metafunction, then E1 and/or E2 may not be true. Common statistical tests can be used to assess these assumptions. If either E1 or E2 are not true, then the OLS estimator is biased and inconsistent. In this case, other consistent estimators are usually available. For example, if E2 is not true, then the explanatory variables are said to be endogenous. In this case, two-stage least squares (2SLS) and generalised method of moments (GMM) estimators are consistent.

Assumptions A7, A8, E1 and E2 are weak by comparison with assumptions that are usually made in the productivity literature. For example, in addition to A7 and A8, growth accountants working in the tradition of Solow (1957) assume that the output distance metafunction is a Cobb-Douglas (CD) function, technical change is Hicks neutral (HN), the technology exhibits constant returns to scale (CRS), markets are perfectly competitive, firms maximise profits, and all variables involved in the production process are observed and measured without error. An important contribution of this paper is to show that if all of these assumptions are true, then the new TFP index takes the form of a GY TFP index.

The structure of the paper is as follows. Section 2 explains in more detail that assumptions A1–A6 guarantee the existence of output and input distance metafunctions.

This section also explains that the value of the output distance metafunction is an output-oriented measure of technical efficiency, and that rational firms may be technically inefficient because they cannot choose technologies freely.

Section 3 explains how output and input distance metafunctions can be used to construct a new proper TFP index. It then explains that common homogeneity, homotheticity and Hicks neutrality assumptions have important implications for the specific form of this index. For example, if technologies are homogeneous of degree r (HDr), then the index can be written in terms of the output distance metafunction only. To improve readability, the proofs of supporting propositions are relegated to the Appendix.

Section 4 defines several measures of efficiency that are of interest to decision-makers. Among them are two measures of efficiency that appear to be new: a measure of environmental efficiency that captures TFP shortfalls associated with changes in the production environment, and a measure of firm efficiency that captures TFP shortfalls associated with variables that are chosen by the firm. The measure of firm efficiency can be exhaustively decomposed into well-known measures of technical, scale and mix efficiency.

Section 5 derives TFP indexes associated with translog, Cobb-Douglas (CD), linear and locally-linear output distance metafunctions. Among other things, this section points out that if assumptions A1–A8 are true, then the output distance metafunction cannot be a translog function.

Section 6 explains that common assumptions concerning markets and firm behaviour have important implications for levels of efficiency, as well as the form of the new TFP index. For example, if output markets are not perfectly competitive, then some firms can maximise revenue by operating inside the production frontier (i.e., by being technically inefficient). This section also shows that if the growth accounting assumptions are true, then the new TFP index takes the form of a GY index.

Section 7 shows that the data envelopment analysis (DEA) models of Banker and Morey (1986), Banker et al. (1984) and Charnes et al. (1978) are all underpinned by assumptions that are rarely, if ever, true. Specifically, these authors implicitly assume that the output distance metafunction is a locally-linear function, and that all variables involved in the production process are observed and measured without error. The fact that these assumptions are rarely true means that most, if not all, DEA estimators are inconsistent estimators of technical efficiency. This does not necessarily mean they are inconsistent estimators of technical efficiency change.

Section 8 explains that if the output distance metafunction exists, then it can be written in the form of a stochastic production frontier. This section also explains that the Battese and Coelli (1995) (hereafter, BC95) inefficiency specification has no information content unless the researcher has made a poor choice of approximating functional form. More seriously, this section points out that the Battese and Coelli (1992) (hereafter, BC92) inefficiency specification is theoretically implausible, so it generally leads to biased and inconsistent estimates of both technical efficiency and technical efficiency change.

Section 9 uses publicly-available U.S. farm data to compute a GY TFP index. Several estimators are then used to decompose the index into measures of technical change and efficiency change. The estimators include a selection of least squares and maximum likelihood estimators. The least squares estimators paint similar pictures of the drivers of TFP change. The maximum likelihood estimators paint qualitatively different pictures of the efficiency change components of productivity change.

Section 10 concludes with a comment on public policy-making and a suggestion for further research.

2. Technologies and Metatechnologies

Consider a production process involving a finite number of variables. In this paper, all variables are classified as either outputs, inputs, or characteristics of the production environment. Outputs and inputs are variables that are *always* chosen by firms, whereas characteristics of the production environment are variables that are *never* chosen by firms. Let $q \in \mathbb{R}_{+}^{N^*}$, $x \in \mathbb{R}_{+}^{M^*}$ and $z \in \mathbb{R}_{+}^{J^*}$ denote vectors of outputs, inputs and environmental variables respectively. The set of output-input combinations that are possible using technology g in a production environment characterised by z is defined as $T(z,g) = \{(x,q)$: inputs x and technology g can produce outputs q in environment $z\}$. The maintained assumptions A1–A6 that were introduced in Section 1 are formally defined as:

A1: $P(x,z,g) \equiv \{q : (x,q) \in T(z,g)\}$ is bounded $\forall x \in \mathbb{R}_{+}^{M^*}$ (output sets bounded),

A2: $q \geq 0 \Rightarrow (0,q) \notin T(z,g)$ (inputs weakly essential), A3: $(x,q) \in T(z,g)$ and $0 \leq \lambda \leq 1 \Rightarrow (x,\lambda q) \in T(z,g)$ (outputs weakly disposable),

A4: $(x,q) \in T(z,g)$ and $\lambda \geq 1 \Rightarrow (\lambda x,q) \in T(z,g)$ (inputs weakly disposable),

A5: $P(x,z,g) \equiv \{q : (x,q) \in T(z,g)\}$ is closed $\forall x \in \mathbb{R}_+^{M^*}$ (output sets closed), and

A6: $L(q,z,g) \equiv \{x : (x,q) \in T(z,g)\}$ is closed $\forall q \in \mathbb{R}_+^{N^*}$ (input sets closed).

If assumptions A1 and A5 are true then the output distance function of Shephard (1970, pp. 207) exists.[3] This function is defined as $D_O(x,q,z,g) = \inf\{\delta > 0 : (x,q/\delta) \in T(z,g)\}$. If A3 is also true then $(x,q) \in T(z,g)$ if and only if $D_O(x,q,z,g) \leq 1$ (Färe and Primont, 1995, p. 22). This means that the production possibilities set and the output distance function are equivalent representations of the technology. On the input side, if A6 is true then the input distance function of Shephard (1970, pp. 206) exists. This function is defined as $D_I(x,q,z,g) = \sup\{\rho > 0 : (x/\rho,q) \in T(z,g)\}$. If A4 is also true then $(x,q) \in T(z,g)$ if and only if $D_I(x,q,z,g) \geq 1$ (Färe and Primont, 1995, p. 22). It follows that if A1–A6 are true then the production possibilities set and the output and input distance functions are equivalent representations of the technology. In this paper, a technology is said to be *regular* if and only if A1–A6 are true.

Technologies that have been developed in the past are generally[4] still available today. In this paper, the set of all technologies available in period t is referred to as the period-t *metatechnology*. If we think of a technology as a book of instructions, then we should think of a metatechnology as a library. In this paper, the period-t metatechnology is said to be regular if and only if all technologies available in period t are regular.

Let Ω_t denote the set of technologies available in period t. The set of output-input combinations that are possible in period t and environment z is $T^t(z) \equiv \cup_{g \in \Omega_t} T(z,g)$. If the metatechnology is regular then equivalent representations of this *metaset* include the output and input distance metafunctions. The output distance metafunction can be defined as $D_O^t(x,q,z) = \min_{g \in \Omega_t} D_O(x,q,z,g) = \inf\{\delta > 0 : (x,q/\delta) \in T^t(z)\}$. The input

[3] Shephard (1970) does not explicitly allow for changes in technologies or production environments. When there is no technical or environmental change, the z and g notation can be suppressed.

[4] Some technologies get lost. For example, Greek fire was a chemical formula that was used in the Byzantine empire for military purposes. The formula was a too-closely-guarded state secret that has since been lost.

distance metafunction can be defined as $D_I^t(x,q,z) = \max_{g \in \Omega_t} D_I(x,q,z,g) = \sup\{\rho > 0 : (x/\rho, q) \in T^t(z)\}$.

The value of the output distance metafunction is a measure of output-oriented technical efficiency (OTE) that is usually attributed to Debreu (1951) and Farrell (1957). In this paper, a firm is said to be technically efficient if $D_O^t(x,q,z) = 1$. As we shall see in Section 6, rational firms may be technically inefficient because they operate in imperfectly competitive markets and are unable to choose inputs and outputs freely. Irrespective of market structure, rational firms may also be technically inefficient because of laws and regulations that prevent them from choosing technologies freely. For example, from 1980–1997, the Cohen-Boyer patents were used by Stanford University to prevent firms from using recombinant-DNA technology without the payment of a license fee. Shortfalls in outputs, revenues and profits due to the choice of technology are sometimes measured using *metatchnology ratios*. For example, output- and input-oriented metatechnology ratios are $OMR(x,q,z,g,t) = D_O^t(x,q,z)/D_O(x,q,z,g)$ and $IMR(x,q,z,g,t) = D_I(x,q,z,g)/D_I^t(x,q,z)$ respectively. If there are no environmental variables involved in the production process and the metatechnology does not change over time, then $OMR(x,q,z,g,t)$ collapses to the metatechnology ratio defined by O'Donnell et al. (2008, p. 236).

3. Regularity Assumptions and Measures of Productivity

In this paper, the TFP of a firm that uses inputs x to produce outputs q is defined as $TFP(x,q) \equiv Q(q)/X(x)$ where $Q(q)$ is an aggregate output, $X(x)$ is a aggregate input, and $Q(.)$ and $X(.)$ are NN, ND and HD1 scalar aggregator functions (O'Donnell, 2012*a*,*d*).[5] If it exists, the output (resp. input) distance metafunction is NN and HD1 in outputs (resp. inputs). If outputs (resp. inputs) are strongly disposable then the output (resp. input) distance metafunction is also ND in outputs (resp. inputs). In this case, natural aggregator functions are $Q(q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})$ and $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$ where $\bar{s}$ is a fixed reference period, and $\bar{x}$, $\bar{q}$, and $\bar{z}$ are fixed vectors of reference inputs, outputs and

[5]The NN, ND and HD1 properties are the distinguishing feature of the O'Donnell (2012*a*,*d*) definition. Defining TFP as the ratio of an aggregate output to an aggregate input is an idea that can be traced back at least as far as Jorgenson and Griliches (1967) and Nadiri (1970). However, Jorgenson and Griliches (1967) are silent on the properties of aggregator functions, and Nadiri (1970, p. 1138, eq. (1)(b)) uses an aggregator function that is not HD1.

environmental variables.[6] The associated measure of TFP is:

$$TFP(x,q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})/D_I^{\bar{s}}(x,\bar{q},\bar{z}). \tag{1}$$

Common regularity assumptions concerning production possibilities metasets can have important implications for output and input distance metafunctions and, therefore, this measure of TFP. Some of the most common assumptions are listed in Table 1.[7] In this table, $a(.)$, $b(.)$, $f(.)$ and $h(.)$ are scalar-valued functions that have properties that are consistent with the properties of $T^t(z)$. For example, if the metatechnology is HDr, then $f(x,z,\bar{x},\bar{z},t)$ is HDr in x and $h(q,z,\bar{q},\bar{z},t)$ is homogeneous of degree $1/r$ in q.

[Table 1 near here]

Assumptions A7 (strong disposability of outputs) and A8 (strong disposability of inputs) were discussed in Section 1. If $Q(q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})$ and A7 is true, then $QI(q,\bar{q}) \equiv Q(q)/Q(\bar{q})$ is a proper output quantity index that compares the output vector q with a reference vector $\bar{q}$. Similarly, if $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$ and A8 is true, then $XI(x,\bar{x}) \equiv X(x)/X(\bar{x})$ is a proper input quantity index that compares the input vector x with a reference vector $\bar{x}$. If $TFP(x,q)$ is given by (1) and A7 and A8 are both true, then $TFPI(x,q,\bar{x},\bar{q}) \equiv TFP(x,q)/TFP(\bar{x},\bar{q})$ is a proper productivity index that compares TFP at the point (x,q) with TFP at the reference point $(\bar{x},\bar{q})$.

Assumption A9 (homogeneity of degree r) says that if inputs are increased by one percent then outputs can be increased by r percent. The metatechnology is said to exhibit decreasing, constant, or increasing returns to scale as r is less than, equal to, or greater than one. If the metatechnology is HDr then $D_O^t(x,q,z) = D_I^t(x,q,z)^{-r}$ (e.g. O'Donnell, 2012*b*, Proposition 6). This means the measure of TFP defined by (1) can be written as $TFP(x,q) = D_O^{\bar{s}}(\bar{x},q,\bar{z}) \times D_O^{\bar{s}}(x,\bar{q},\bar{z})^{1/r}$ (i.e., in terms of the output distance metafunction only). The practical significance of this result is that we only need to estimate the output distance metafunction in order to estimate TFP change.

[6] Balk (1998, p. 19) also uses the overbar notation to denote reference points. However, in statistics, overbars are commonly used to denote sample means. Any confusion about this notation will, at worst, lead researchers to use sample means as points of reference. In economics, unit vectors are also used as reference vectors (e.g., Balk, 1998, p. 16, fn. 4).

[7] For proofs of D9–D13, see O'Donnell (2012*b*, Propositions 3 and 4) and Propositions 1–4 in the Appendix of this paper.

Assumption A10 (implicit Hicks output neutrality) says that the outputs that can be produced using given inputs in a given environment in a given period are a scalar multiple of the outputs that can be produced using the same inputs in the same environment in any other period. If technical change is implicit Hicks output neutral (IHON) then marginal rates of technical transformation are time-invariant. On the other hand, assumption A11 (output homotheticity) says that the outputs that can be produced in a given period using given inputs in a given environment are a scalar multiple of the outputs that can be produced in the same period using any other inputs in any other environment. If a metatechnology is output homothetic (OH) then marginal rates of technical substitution do not depend on outputs and marginal rates of technical transformation do not depend on inputs. If technical change is IHON and the metatechnology is OH then the output distance metafunction takes the form $D_O^t(x,q,z) = Q(q)/F^t(x,z)$ where $Q(q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})$ is an aggregate output and $F^t(x,z)$ is the maximum aggregate output possible in period t using input vector x in an environment characterised by z (i.e., it is a production function) (Proposition 5). Thus, we can write $\ln Q(q) = \ln F^t(x,z) - u$ where $u \equiv -\ln D_O^t(x,q,z) \geq 0$ is an output-oriented technical inefficiency effect. This equation is the bedrock of the production frontier literature. The measure of TFP defined by (1) can also be written as $TFP(x,q) = [F^t(x,z)/X(x)]\exp(-u)$ where $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$. This equation provides a basis for an output-oriented decomposition of productivity change.

Assumption A12 (implicit Hicks input neutrality) says that the inputs that can produce a given output vector in a given environment in a given period are a scalar multiple of the inputs that can produce the same outputs in the same environment in any other period. If technical change is implicit Hicks input neutral (IHIN) then marginal rates of technical substitution are time-invariant. Similarly, assumption A13 (input homotheticity) says that the inputs that can produce a given output vector in a given environment in a given period are a scalar multiple of the inputs that can produce any other outputs in any other environment in that period. If a metatechnology is input homothetic (IH) then, again, marginal rates of technical substitution do not depend on outputs and marginal rates of technical transformation do not depend on inputs. If technical change is IHIN and the metatechnology is IH then the input distance metafunction takes the form $D_I^t(x,q,z) = X(x)/H^t(q,z)$ where $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$ is an aggregate input and $H^t(q,z)$ is the minimum aggregate input capable of producing q in period t in an environment characterised by z (i.e., it is an input function) (Proposition 6). In this case, we can write $-\ln X(x) = -\ln H^t(q,z) - u^*$ where $u^* \equiv \ln D_I^t(x,q,z) \geq 0$ is an input-oriented technical inefficiency effect. This equation underpins the stochastic frontier model of O'Donnell and Nguyen (2013, eq. 21). The measure of TFP defined by (1) can also be written as $TFP(x,q) = [Q(q)/H^t(q,z)]\exp(-u^*)$ where $Q(q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})$. This equation provides a basis for an input-oriented decomposition of productivity change.

Finally, in this paper, if technical change is both IHON and IHIN then and only then it is said to be Hicks neutral (HN). Similarly, if a metatechnology is both OH and IH then and only then it is said to be homothetic (H).[8] If technical change is HN and the metatechnology is H and HDr, then $D_O^t(x,q,z) = Q(q)/[A^t(z)X(x)^r]$ where $Q(q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})$ is an aggregate output, $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$ is an aggregate input, and $A^t(z) \propto 1/D_O^t(\bar{x},\bar{q},z)$ is the Solow (1957) residual (Proposition 7).[9] In this case, the measure of TFP defined by (1) can be written as $TFP(x,q) = A^t(z)X(x)^{r-1}D_O^t(x,q,z)$. If firms are technically efficient and the metatechnology exhibits CRS then the last two terms vanish and $TFP(x,q) = A^t(z)$ (i.e., the Solow residual can be interpreted as a measure of TFP).

4. Measures of Efficiency

It is convenient at this point to introduce firm and time subscripts into the notation. Let $x_{it} = (x_{1it},\ldots,x_{Mit})'$ and $q_{it} = (q_{1it},\ldots,q_{Nit})'$ denote the inputs and outputs of firm i in period t, and let $TFP_{it} \equiv TFP(x_{it},q_{it})$ denote the associated measure of TFP. Decision-makers are often interested in the relationship between TFP and various measures of efficiency. Measures of efficiency that are of particular interest include TFP efficiency (TFPE), firm efficiency (FE), environmental efficiency (EE), output-oriented technical efficiency (OTE) and output-oriented scale-mix efficiency (OSME). In this

[8]Elsewhere in the production economics literature, if a technology is both OH and IH then it is said to be *inversely* homothetic (e.g., Färe and Primont, 1995, p. 70). This paper uses the term homothetic to avoid any confusion with the use of the acronym IH.

[9]Solow (1957, p. 312) uses t to represent "*any kind of shift* in the production function" (his italics) and writes $A(t)$ instead of $A^t(z)$.

paper, the TFPE, FE, EE, OTE and OSME of firm i in period t are formally defined as:

$$TFPE_{it} \equiv TFP_{it}/TFP_t^*, \tag{2}$$

$$FE_{it} \equiv TFP_{it}/TFP^t(z_{it}), \tag{3}$$

$$EE_{it} \equiv TFP^t(z_{it})/TFP_t^*, \tag{4}$$

$$OTE_{it} \equiv D_O^t(x_{it}, q_{it}, z_{it}) \tag{5}$$

and

$$OSME_{it} \equiv FE_{it}/OTE_{it} \tag{6}$$

where TFP_t^* denotes the maximum TFP that is possible in period t, and $TFP^t(z_{it})$ is the maximum TFP that is possible in period t in an environment characterised by z_{it}. TFPE is an overall measure of productive performance defined by O'Donnell (2012*d*, p. 880). As the name implies, FE is a measure of efficiency associated with variables that are, at some point, chosen by firms (i.e., inputs, outputs and technologies). Similarly, EE is a measure of efficiency associated with variables that are *never* chosen by firms (i.e., characteristics of the production environment). These two measures appear to be new. If there are no environmental variables involved in the production process then $EE_{it} = 1$ and $TFPE_{it} = FE_{it}$ for all i and t. Finally, OTE is the Debreu-Farrell output-oriented measure of technical efficiency discussed in Section 2, and OSME is the measure of economies of scale and scope defined by O'Donnell (2012*a*). It is clear from equations (2) to (6) that $TFPE_{it} = FE_{it} \times EE_{it}$ and $FE_{it} = OTE_{it} \times OSME_{it}$. Thus, the TFPE of firm i in period t can be written as $TFPE_{it} = EE_{it} \times OTE_{it} \times OSME_{it}$. This provides a basis for an exact output-oriented decomposition of any proper TFP index.

5. Functional Forms and the Components of Productivity Change

The index that compares the TFP of firm i in period t with the TFP of firm k in period s is defined as $TFPI_{ksit} \equiv TFP_{it}/TFP_{ks}$. If $Q(q) = D_O^{\bar{s}}(\bar{x}, q, \bar{z})$ and $X(x) = D_I^{\bar{s}}(x, \bar{q}, \bar{z})$ then

$$TFPI_{ksit} = \frac{D_O^{\bar{s}}(\bar{x}, q_{it}, \bar{z})}{D_O^{\bar{s}}(\bar{x}, q_{ks}, \bar{z})} \frac{D_I^{\bar{s}}(x_{ks}, \bar{q}, \bar{z})}{D_I^{\bar{s}}(x_{it}, \bar{q}, \bar{z})}. \tag{7}$$

If there are no environmental variables involved in the production process then this new index collapses to the Färe-Primont TFP index defined by O'Donnell (2012*c*).

The regularity assumptions discussed in Section 3 have obvious implications for the mathematical form of (7). For example, if technical change is HN and the metatechnology is H and HDr, then (7) can be written as

$$TFPI_{ksit} = \left[\frac{A^t(z_{it})}{A^s(z_{ks})}\right]\left[\frac{D_O^t(x_{it},q_{it},z_{it})}{D_O^s(x_{ks},q_{ks},z_{ks})}\right]\left[\frac{X(x_{it})}{X(x_{ks})}\right]^{r-1} \tag{8}$$

where $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$ and $A^t(z) \propto 1/D_O^t(\bar{x},\bar{q},z)$. If there are no environmental variables involved in the production process then this equation collapses to equation (19) in O'Donnell (2012*c*). The first term on the right-hand side of (8) is a combined measure of technical change (hereafter abbreviated as *d*T) and environmental efficiency change (*d*EE), the second term is a measure of output-oriented technical efficiency change (*d*OTE), and the last term is a measure of output-oriented scale efficiency change (*d*OSE). Thus, equation (8) says that the TFP index (7) can be decomposed as *d*TFP $=$ *d*T $\times$ *d*EE $\times$ *d*OTE $\times$ *d*OSE. The HN and H assumptions mean that the aggregator functions $Q(q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})$ and $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$ are proportional to the output and input distance metafunctions. This means that all input and output mixes are equally productive and there are no mix efficiency change components in (8). For details, see O'Donnell (2012*a*, Section 3.7).

In practice, it is common to assume that the output distance metafunction is a translog, CD, linear or locally-linear function. To assume it is a translog function is to assume it is a linearly homogeneous (in outputs) version of the translog function of Heady and Dillon (1961, p. 205) and Christensen et al. (1973):

$$\begin{aligned}
\ln D_O^t(x_{it},q_{it},z_{it}) = {} & \sum_{n=1}^{N^*}\alpha_{nt}\ln q_{nit} - \gamma_t - \sum_{j=1}^{J^*}\rho_{jt}\ln z_{jit} - \sum_{m=1}^{M^*}\beta_{mt}\ln x_{mit} \\
& - \sum_{j=1}^{J^*}\sum_{k=1}^{J^*}\rho_{jkt}\ln z_{jit}\ln z_{kit} - \sum_{m=1}^{M^*}\sum_{j=1}^{J^*}\phi_{mjt}\ln x_{mit}\ln z_{jit} \\
& - \sum_{n=2}^{N^*}\sum_{j=1}^{J^*}\theta_{njt}\ln(q_{nit}/q_{1it})\ln z_{jit} - \sum_{m=1}^{M^*}\sum_{k=1}^{M^*}\beta_{mkt}\ln x_{mit}\ln x_{kit} \\
& - \sum_{n=2}^{N^*}\sum_{m=1}^{M^*}\delta_{nmt}\ln(q_{nit}/q_{1it})\ln x_{mit} \\
& - \sum_{n=2}^{N^*}\sum_{k=2}^{N^*}\alpha_{nkt}\ln(q_{nit}/q_{1it})\ln(q_{kit}/q_{1it})
\end{aligned} \tag{9}$$

where $\rho_{jkt} = \rho_{kjt}$, $\phi_{mjt} = \phi_{jmt}$, $\theta_{njt} = \theta_{jnt}$, $\beta_{mkt} = \beta_{kmt}$, $\delta_{nmt} = \delta_{mnt}$, $\alpha_{nkt} = \alpha_{knt}$ and $\sum_n \alpha_{nit} = 1$. This function cannot satisfy *any* of assumptions A7–A13. The fact that it cannot satisfy A7 or A8 (the strong disposability properties) means it cannot be used to construct proper quantity indexes.[10] The fact that it cannot satisfy A9 (HDr) means that it cannot be used on its own to construct the TFP index (7) (we also need the associated input distance metafunction). The fact that it cannot satisfy A10 or A12 (the IHON and IHIN properties) simply means that the researcher must select $\bar{s}$ in order to construct (7). Outputs will be strongly disposable (SD) if $\theta_{njt} = \delta_{nmt} = \alpha_{nkt} = 0$ and $\alpha_{nt} \geq 0$, inputs will be SD if $\phi_{mjt} = \delta_{nmt} = \beta_{mkt} = 0$ and $\beta_{mt} \geq 0$, the HDr property will be satisfied if $\sum_n \delta_{nmt} = \sum_j \phi_{jmt} = \sum_k \beta_{kmt} = 0$ and $\sum_m \beta_{mt} = r$, and technical change will be HN if $\alpha_{nt} = \alpha_n$ and $\beta_{mt} = \beta_m$. If all of these SD, HDr and HN constraints are satisfied, then $D_O^{\bar{s}}(\bar{x}, q_{it}, \bar{z}) \propto \prod_n q_{nit}^{\alpha_n}$ and $D_I^{\bar{s}}(x_{it}, \bar{q}, \bar{z}) \propto \prod_m x_{mit}^{\lambda_m}$ where $\lambda_m \equiv \beta_m / r$. In this case, the TFP index (7) takes the form:

$$TFPI_{ksit} = \prod_{n=1}^{N^*} \left(\frac{q_{nit}}{q_{nks}} \right)^{\alpha_n} \prod_{m=1}^{M^*} \left(\frac{x_{mks}}{x_{mit}} \right)^{\lambda_m}. \tag{10}$$

O'Donnell (2012*b*) refers to these types of indexes as *multiplicative* indexes. Imposing the SD, HDr and HN constraints means that the metafunction will also satisfy A11 and A13 (the OH and IH properties). Thus, this multiplicative index can also be written in the form of (8).

The log-linear function of Cobb and Douglas (1928) is a special case of (9) with all second-order coefficients constrained to zero. In the CD case, the SD, HDr and HN properties will be satisfied if $\alpha_{nt} \geq 0$, $\beta_{mt} \geq 0$, $\sum_m \beta_{mt} = r$, $\alpha_{nt} = \alpha_n$ and $\beta_{mt} = \beta_m$. It almost goes without saying that if these constraints are satisfied then the TFP index (7) can be written in the form of (8) or (10).

If the output distance metafunction is a linear function then it takes the form

$$D_O^t(x_{it}, q_{it}, z_{it}) = \alpha_t' q_{it} / (\gamma_t + \rho_t' z_{it} + \beta_t' x_{it}). \tag{11}$$

The term *linear* derives from the fact that if $D_O^t(x_{it}, q_{it}, z_{it}) = 1$ then $\alpha_t' q_{it} = \gamma_t - \rho_t' z_{it} -$

[10] Translog functions are often used to motivate the use of Törnqvist indexes. Törnqvist indexes are intransitive, so they are not proper indexes for multitemporal or multilateral comparisons. However, they are still proper indexes for binary comparisons. For details, see O'Donnell (2012*b*).

$\beta_t' x_{it}$ (i.e., the frontier is linear in the output, input and environmental variables). The SD properties will be satisfied if $\alpha_t \geq 0$ and $\beta_t \geq 0$, the HDr property cannot be satisfied for $r \neq 1$ but will be satisfied for $r = 1$ if $\gamma_t + \rho_t' z_{it} = 0$, and the HN property will be satisfied if $\alpha_t = \alpha$ and $\beta_t = \beta$. If all of these constraints are satisfied then $D_O^{\bar{s}}(\bar{x}, q_{it}, \bar{z}) \propto \alpha' q_{it}$ and $D_I^{\bar{s}}(x_{it}, \bar{q}, \bar{z}) \propto \beta' x_{it}$. In this case, the TFP index (7) takes the form:

$$TFPI_{ksit} = \left(\frac{\alpha' q_{it}}{\alpha' q_{ks}}\right)\left(\frac{\beta' x_{ks}}{\beta' x_{it}}\right). \tag{12}$$

Balk (2008, p. 65) refers to these types of indexes as *additive* indexes. Imposing the HD1 and HN constraints also means the metafunction is homothetic (H). Thus, the TFP index (12) can also be written in the form of (8). However, the combination of linearity, SD, HD1 and HN means there are no technical, environmental or scale efficiency change components in (8) (i.e., linearity, SD, HD1 and HN $\Rightarrow$ dTFP $=$ dOTE).

Finally, if the output distance metafunction is a locally-linear function then

$$D_O^t(x_{it}, q_{it}, z_{it}) = \alpha_{it}' q_{it} / (\gamma_{it} + \rho_{it}' z_{it} + \beta_{it}' x_{it}). \tag{13}$$

As we shall see in Section 7, this assumption underpins the DEA approach to measuring OTE. The term *local* derives from the fact that, unlike the parameters in (11), the parameters in (13) vary across both firms and time periods (i.e., there is a potentially different linear relationship between the variables in the neighbourhood of each data point). Again, the SD properties will be satisfied if $\alpha_{it} \geq 0$ and $\beta_{it} \geq 0$, the HDr property cannot be satisfied for $r \neq 1$ but will be satisfied for $r = 1$ if $\gamma_{it} + \rho_{it}' z_{it} = 0$, and the HN property will be satisfied if $\alpha_{it} = \alpha$ and $\beta_{it} = \beta$. If the HN constraint is satisfied then (13) is still linear, but no longer locally linear. Thus, the combination of local-linearity, SD, HD1 and HN means that the TFP index (7) can again be written in the form of (8) or (12). However, once again, there are no technical, environmental or scale efficiency change components in (8) (i.e., dTFP $=$ dOTE).

6. Markets, Behaviour and Technical Efficiency

The output distance metafunction is a representation of the set of output-input combinations that are *technically feasible*. Beyond this, the metafunction provides no information about the output-input combinations that rational firms might *choose*. For example, it does not explain why some firms can maximise revenue by operating inside the production frontier (i.e., by being technically inefficient).[11] These outcomes have a lot to do with market structure. In the productivity literature, it is common to assume that output and/or input markets are perfectly competitive, and that firms maximise revenue and/or minimise cost.[12]

If output markets are perfectly competitive then output prices do not depend on the outputs of any single firm (i.e., firms are *price-takers*). Let d_{it} denote a vector of variables that can affect demand for the outputs of firm i in period t (e.g., if the firm produces intermediate goods then these so-called *demand shifters* may include characteristics of downstream production environments; if the firm produces final goods then they may include consumer incomes and preferences). In general, the inverse demand correspondence for firm i in period t is a function of the form $p(q_{it}, d_{it}) \geq 0$. If output markets are perfectly competitive then $\partial p_n(q_{it}, d_{it})/\partial q_{kit} = 0$ for $n = 1, \ldots, N$ and $k = 1, \ldots, N$, where $p_n(q_{it}, d_{it})$ denotes the n-th element of $p(q_{it}, d_{it})$.

To assume that firms maximise revenue is to assume that every firm in the population actually maximises revenue. The period-t revenue maximisation problem of firm i is

$$R^t(x_{it}, d_{it}, z_{it}) = \max_{q \geq 0} \{p(q, d_{it})'q : D_O^t(x_{it}, q, z_{it}) \leq 1\}. \tag{14}$$

Let $q_{it}^* = q^t(x_{it}, d_{it}, z_{it})$ denote the output vector that solves (14). To assume that firms maximise revenue is to assume that $q_{it} = q_{it}^*$ for all i and t. Importantly, the value of the output distance metafunction (i.e., the level of OTE) depends on the structure of, and possibly conditions in, downstream markets. To see this, it is useful to write the n-th first-order condition (FOC) for a maximum as $\partial p(q, d_{it})/\partial \ln q_n + p_n(q, d_{it}) - \lambda \partial D_O^t(x_{it}, q, z_{it})/\partial q_n = 0$. The complementary slackness condition is $\lambda [D_O^t(x_{it}, q, z_{it}) - 1] = 0$. If output markets are perfectly competitive then $\partial p(q, d_{it})/\partial \ln q_n = 0$, the optimal solution for the Kuhn-Tucker multiplier is $\lambda_{it}^* > 0$, and the revenue maximising output vector satisfies $D_O^t(x_{it}, q_{it}^*, z_{it}) = 1$. Thus, to assume that firms maximise revenue and that output markets are perfectly competitive is to assume that $OTE_{it} \equiv D_O^t(x_{it}, q_{it}, z_{it}) =$

[11] Leibenstein (1979) refers to this type of behaviour as *X-inefficiency*.

[12] For example, these assumptions underpin the growth accounting approach to measuring productivity change, where they are used to motivate the computation of Törnqvist productivity indexes.

1. Of course, if output markets are *not* perfectly competitive then $\partial p(q, d_{it})/\partial \ln q_n \neq 0$ and $\lambda_{it}^* = 0$ is possible. If $\lambda_{it}^* = 0$ then the complementary slackness condition implies $D_O^t(x_{it}, q_{it}^*, z_{it}) < 1$. Thus, to assume that firms maximise revenue and that output markets are *not* perfectly competitive is to assume that $OTE_{it} \leq 1$. For policy-makers, the most important implication of this result is that if output markets are not perfectly competitive then firms can be both revenue efficient and technically inefficient at the same time.[13] For econometricians, the most important implication is that if output markets are not perfectly competitive then[14]

$$u_{it} \equiv -\ln D_O^t(x_{it}, q_{it}, z_{it}) = U^t(x_{it}, d_{it}, z_{it}) \geq 0. \tag{15}$$

This variable is the inefficiency effect in the SFA models of Aigner et al. (1977) and Meeusen and van den Broeck (1977). Equation (15) says that this effect can be written as a function of variables that are not directly involved in the production process (in this case, demand shifters).

To make these ideas more concrete, consider a single-output firm and suppose the inverse demand function is an additively separable function of the form $p(q_{it}, d_{it}) = p(d_{it}) - bq_{it}$, where $p(.)$ is a nonnegative function and $b \geq 0$. In this case, the n-th FOC for a maximum is $p(d_{it}) - 2bq - \lambda = 0$. Moreover, the complementary slackness condition is $\lambda[q - F^t(x_{it}, z_{it})] = 0$. It is easily shown that:[15]

$$q_{it}^* = \begin{cases} p(d_{it})/2b & \text{if } p(d_{it}) < 2bF^t(x_{it}, z_{it}), \\ F^t(x_{it}, z_{it}) & \text{otherwise.} \end{cases} \tag{16}$$

[13] Mathematically, revenue efficiency is defined as $RE_{it} \equiv p_{it}' q_{it}/R^t(x_{it}, d_{it}, z_{it})$. The firm is revenue efficient if and only if $RE_{it} = 1$.

[14] Revenue maximisation means $q_{it} = q_{it}^* = q^t(x_{it}, d_{it}, z_{it})$. Substituting this into $D_O^t(x_{it}, q_{it}, z_{it})$ yields a function that depends on (x_{it}, d_{it}, z_{it}) instead of (x_{it}, q_{it}, z_{it}).

[15] There are three possible solutions to the complementary slackness condition. If $q < F^t(x_{it}, z_{it})$ and $\lambda = 0$ and then the FOC $\Rightarrow q = p(d_{it})/2b < F^t(x_{it}, z_{it})$. If $q = F^t(x_{it}, z_{it})$ and $\lambda = 0$ then the FOC $\Rightarrow q = p(d_{it})/2b = F^t(x_{it}, z_{it})$. If $q = F^t(x_{it}, z_{it})$ and $\lambda > 0$ then the FOC $\Rightarrow \lambda = p(d_{it}) - 2bF^t(x_{it}, z_{it}) > 0 \Rightarrow p(d_{it})/2b > F^t(x_{it}, z_{it}) = q$.

If firms maximise revenue, then $q_{it} = q^*_{it}$ and

$$u_{it} = \begin{cases} \ln(2b) + \ln F^t(x_{it}, z_{it}) - \ln p(d_{it}) & \text{if } p(d_{it}) < 2bF^t(x_{it}, z_{it}), \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

In this simple example, if output markets are perfectly competitive, then $b = 0$ and $u_{it} = 0$ (i.e., the firm is fully efficient). If output markets are not perfectly competitive and demand conditions weaken to the point where $p(d_{it})/2b < F^t(x_{it}, z_{it})$, then the firm will choose to operate below the frontier (i.e., at less than full capacity). If demand in the output market is sufficiently strong, then the firm will operate on the frontier (i.e., at full capacity).

Analogous results are available for the case where firms minimise cost. On the input side, the inverse supply correspondence is a function of the form $w(x_{it}, s_{it}) \geq 0$ where s_{it} is a vector of variables that can affect the supply of inputs (e.g., characteristics of upstream production environments). If input markets are perfectly competitive then input prices do not depend on the amounts purchased by any single firm. Mathematically, this means $\partial w_m(x_{it}, s_{it})/\partial x_{kit} = 0$ for $m = 1, \ldots, M$ and $k = 1, \ldots, M$, where $w_m(x_{it}, s_{it})$ denotes the m-th element of $w(x_{it}, s_{it})$. It is easily shown that if firms minimise cost and input markets are perfectly competitive then $OTE_{it} = 1$. If firms minimise cost and input markets are *not* perfectly competitive then $OTE_{it} \leq 1$. Thus, if input markets are not perfectly competitive then firms can be cost efficient and technically inefficient at the same time. Moreover, the inefficiency effect in the SFA model can again be written as a function of variables that are not directly involved in the production process (in this case, supply shifters).

Finally, it is evident that different assumptions concerning markets and behaviour can have important implications for measures of technical efficiency. By extension, these assumptions can have important implications for the technical efficiency change component of TFP change, and possibly the form of the TFP index (7). To see this more clearly, and with a view to the empirical illustration in Section 9, suppose that output and input markets are perfectly competitive, firms maximise profit (implying they also maximise revenue and minimise cost), technical change is HN, and the output distance metafunction is a CD function. These assumptions underpin the growth accounting approach to measuring productivity change. In this case, profit-maximising revenue and cost shares are $r^*_{nit} = \alpha_n$ and $s^*_{mit} = \lambda_m$. If firms maximise profit then $r_{nit} = \alpha_n = \bar{r}_n$ and $s_{mit} = \lambda_m = \bar{s}_m$ where $\bar{r}_n$ and $\bar{s}_m$ are the sample mean revenue and cost shares respectively. Thus, the TFP index (7) collapses to the GY TFP index defined by O'Donnell (2011, p. 43):[16]

$$TFPI^{GY}_{ksit} = \left[\prod_{n=1}^{N^*}\left(\frac{q_{nit}}{q_{nks}}\right)^{\bar{r}_n}\right]\left[\prod_{m=1}^{M^*}\left(\frac{x_{mks}}{x_{mit}}\right)^{\bar{s}_m}\right]. \tag{18}$$

Like all proper TFP indexes, the this index can be decomposed into a measure of technical change and various measures of efficiency change. For example, if technical change is HN and the output distance metafunction is a CD function, then (18) can be decomposed as:

$$\begin{aligned} TFPI^{GY}_{ksit} = {} & \left[\frac{\exp(\gamma_1 t)}{\exp(\gamma_1 s)}\right]\prod_{j=1}^{J^*}\left(\frac{z_{jit}}{z_{jks}}\right)^{\rho_j}\left[\frac{D^t_O(x_{it}, q_{it}, z_{it})}{D^s_O(x_{ks}, q_{ks}, z_{ks})}\right] \\ & \times \prod_{n=1}^{N^*}\left(\frac{q_{nit}}{q_{nks}}\right)^{\kappa_n}\prod_{m=1}^{M^*}\left(\frac{x_{mit}}{x_{mks}}\right)^{\tau_m} \end{aligned} \tag{19}$$

where $\kappa_n \equiv \bar{r}_n - \alpha_n$ and $\tau_m \equiv \beta_m - \bar{s}_m$. The first term is a measure of technical change, the second term is a measure of EE change, the third term is a measure of OTE change, and the remaining terms (on the second line) collectively measure OSME change (hereafter dOSME). Thus, equation (19) says that $d\text{TFP} = d\text{T} \times d\text{EE} \times d\text{OTE} \times d\text{OSME}$. If all of the growth accounting assumptions are true, then $d\text{OTE} \times d\text{OSME} = 1$. TFP indexes associated with locally-linear metafunctions can also be exhaustively decomposed in this way (e.g., O'Donnell, 2012*d*).

7. Data Envelopment Analysis

To estimate measures of technical efficiency and productivity change, it is common to assume that outputs and inputs are strongly disposable and that the metafunction is locally linear. DEA models are also underpinned by the assumption that all variables involved in the production process are observed and measured without error.

Suppose that observations on I firms over t periods are available for estimating the

[16]The first term in square brackets is the GY output quantity index defined by O'Donnell (2012*b*, eq. 5).

period-t metatechnology. If the metafunction is given by (13) (i.e., locally-linear), then estimating the OTE of firm i in period t involves choosing the unknown parameters α_{it}, γ_{it}, ρ_{it} and β_{it} to maximise $D^t_O(x_{it}, q_{it}, z_{it}) = \alpha'_{it} q_{it}/(\gamma_{it} + \rho'_{it} z_{it} + \beta'_{it} x_{it}) \leq 1$. The chosen parameters must satisfy two sets of constraints. First, if the chosen parameters are used to evaluate the metafunction at *any* data point in the sample, then the value obtained must be no more than one (equivalently, if all the data are depicted on a scatter diagram, every hyperplane that forms a part of the estimated frontier must envelop all the points in the scatter). The second set of constraints are the SD constraints discussed in Section 5, namely $\alpha_{it} \geq 0$ and $\beta_{it} \geq 0$. Thus, the researcher's optimisation problem takes the form:

$$\widehat{D^t_O}(x_{it}, q_{it}, z_{it}) = \max_{\alpha_{it}, \gamma_{it}, \rho_{it}, \beta_{it}} \quad \alpha'_{it} q_{it}/(\gamma_{it} + \rho'_{it} z_{it} + \beta'_{it} x_{it}) \tag{20a}$$

$$\text{s.t.} \quad \alpha'_{it} q_{ks}/(\gamma_{it} + \rho'_{it} z_{ks} + \beta'_{it} x_{ks}) \leq 1 \text{ for all } k \text{ and } s, \tag{20b}$$

$$\text{and} \quad \alpha_{it}, \beta_{it} \geq 0. \tag{20c}$$

Unfortunately, this constrained maximisation problem has an infinite number of solutions. A unique solution is usually identified by setting $\alpha'_{it} q_{it} = 1$. With this normalisation, the researcher's optimization problem can be written as:

$$\widehat{D^t_O}(x_{it}, q_{it}, z_{it})^{-1} = \min_{\alpha_{it}, \gamma_{it}, \rho_{it}, \beta_{it}} \quad \gamma_{it} + \rho'_{it} z_{it} + \beta'_{it} x_{it} \tag{21a}$$

$$\text{s.t.} \quad \gamma_{it} + \rho'_{it} z_{ks} + \beta'_{it} x_{ks} - \alpha'_{it} q_{ks} \geq 0 \text{ for all } k \text{ and } s, \tag{21b}$$

$$\alpha'_{it} q_{it} = 1, \tag{21c}$$

$$\text{and} \quad \alpha_{it}, \beta_{it} \geq 0. \tag{21d}$$

This is the DEA estimator of the reciprocal of OTE under the assumption of variable returns to scale (VRS). Imposing the restriction $\gamma_{it} + \rho'_{it} z_{it} = 0$ will ensure the estimated metatechnology exhibits CRS; the restriction $\gamma_{it} + \rho'_{it} z_{it} \leq 0$ will ensure it exhibits nondecreasing returns to scale (NDRS); and the restriction $\gamma_{it} + \rho'_{it} z_{it} \geq 0$ will ensure it exhibits nonincreasing returns to scale (NIRS). Non-sample information concerning the effects of changes in the production environment can be accommodated by adding sign restrictions on the elements of ρ_{it}.

LP (21) is a *primal* output-oriented DEA LP. Every primal LP has a *dual* form with the property that if the primal and the dual LPs both have feasible solutions then the optimised values of the two objective functions are equal. The dual form of LP (21) is:

$$
\begin{aligned}
\widehat{D_O^t}(x_{it}, q_{it}, z_{it})^{-1} = \max_{\mu,\lambda} \quad & \mu && \text{(22a)}\\
\text{s.t.} \quad & \mu q_{nit} - \sum_{k=1}^{I}\sum_{s=1}^{t} \lambda_{ks} q_{nks} \leq 0 \ \text{ for } \ n = 1\ldots, N, && \text{(22b)}\\
& \sum_{k=1}^{I}\sum_{s=1}^{t} \lambda_{ks} x_{mks} \leq x_{mit} \ \text{ for } \ m = 1\ldots, M, && \text{(22c)}\\
& \sum_{k=1}^{I}\sum_{s=1}^{t} \lambda_{ks} z_{jks} = z_{jit} \ \text{ for } \ j = 1\ldots, J, && \text{(22d)}\\
& \lambda' \iota = 1 && \text{(22e)}\\
\text{and} \quad & \lambda, \mu \geq 0 && \text{(22f)}
\end{aligned}
$$

where $\lambda = (\lambda_{11}, \ldots, \lambda_{It})'$ and $\iota = (1, \ldots, 1)'$. This LP seeks to scale up the output vector while holding the input vector and characteristics of the production environment fixed. The role of the constraints (22b)–(22f) is to define the boundary of the production possibilities set. Sign restrictions on the elements of ρ_{it} in the primal LP (21) will affect the equality signs in the dual constraint (22d). For example, if the j-th element of ρ_{it} is constrained to be nonpositive then the equality sign in the j-th constraint becomes a "$\geq$" sign. If there are no environmental variables involved in the production process then (22d) is deleted from the problem altogether. In this case, the LP reduces to an output-oriented version of an LP that can be traced back to Banker et al. (1984, p. 1084, eq. 19). If the metatechnology exhibits CRS then (22d) and (22e) are replaced with

$$
\sum_{k=1}^{I}\sum_{s=1}^{t} (\lambda_{ks} + \kappa_{ks}) z_{jks} = z_{jit} \ \text{ for } \ j = 1, \ldots, J \tag{23}
$$

$$
\text{and} \quad \sum_{k=1}^{I}\sum_{s=1}^{t} (\lambda_{ks} + \kappa_{ks}) = 1 \tag{24}
$$

where $\kappa_{11}, \ldots, \kappa_{It}$ are unsigned. If there are no environmental variables and the metatechnology exhibits CRS then (22d) and (22e) are both deleted. In this case, the LP reduces to an output-oriented version of an LP that is usually attributed to Afriat (1972) and Charnes et al. (1978). If the metatechnology exhibits NIRS then (22d) and (22e) are still replaced with (23) and (24), but $\kappa_{11},\ldots,\kappa_{It}$ are constrained to be nonnegative. Finally, if there are no environmental variables and the technology exhibits NIRS then (22d) is deleted from the problem and (22e) is replaced with $\lambda'\iota \leq 1$.

The different LPs described above are, in fact, different DEA estimators of (the reciprocal of) OTE. These estimators are biased in finite samples. However, if the assumptions underpinning the relevant LPs are true, they are consistent estimators.[17] Among other things, this means that consistent estimates of technical, scale and mix efficiency can be obtained using LPs of the type described in O'Donnell (2012*d*). Unfortunately, if the assumptions underpinning the relevant LPs are *not* true, then the associated DEA estimators are inconsistent. In practice, the assumptions underpinning DEA LPs are rarely, if ever, true (e.g., output, input and environmental variables are almost always measured with error, if not unobserved). It follows that most, if not all, DEA estimators are inconsistent. This has important practical implications: as Simar and Wilson (2000, pp. 56, 77) put it, "Consistency is an essential property for any estimator. Indeed, it would be rather meaningless to use an estimator that does not satisfy consistency, since even with an infinite amount of data, an inconsistent estimator cannot be expected to give an accurate estimate of the quantity of interest ... If the data contain noise, DEA and FDH estimators will be inconsistent, and there seems little choice but to rely on SFA."

8. Stochastic Frontier Analysis

In most practical situations, the functional form of the output distance metafunction is unknown. Furthermore, not all variables involved in the production process are observed and measured without error. Without loss of generality, suppose the first M inputs and the first J environmental variables are observed. If the output distance metafunction

[17] The prospect of getting more reliable estimates of OTE by increasing the number of observations in the sample has led to interest in DEA bootstrapping methods. Relevant literature on bias, consistency and bootstrapping can be accessed from Kneip et al. (2011).

exists then it is a fact (not an assumption) that

$$\ln Q_{it} = \gamma_t + f(x_{1it}, \ldots, x_{Mit}, z_{1it}, \ldots, z_{Jit}) + v_{it} - u_{it} \tag{25}$$

where Q_{it} is an observed aggregate output, γ_t is an unknown parameter, $f(.)$ is a known approximating function that is linear in the parameters, $u_{it} \equiv -\ln D_O^t(x_{it}, q_{it}, z_{it}) \geq 0$ is an unobserved inefficiency effect, and $v_{it} = \ln Q_{it} - \gamma_t - f(x_{1it}, \ldots, x_{Mit}, z_{1it}, \ldots, z_{Jit}) - \ln D_O^t(x_{it}, q_{it}, z_{it})$ is an unobserved variable accounting for statistical noise. Statistical noise comprises functional form errors (e.g., the possibility that the unknown output distance metafunction is not, in fact, separable in outputs), omitted variable errors (e.g., the possibility that $M < M^*$) and measurement errors (e.g., failing to account for quality by classifying goods and services into homogeneous groups). If any inputs, outputs or environmental variables are random, then this noise component will generally also be random. In this case, frontier production is also random and (25) is a panel data version of the stochastic production frontier of Aigner et al. (1977). If inputs and environmental variables are not random and there is no statistical noise, then (25) is a deterministic production frontier of the type considered by Aigner and Chu (1968).

Characteristics of technologies, markets and behaviour have important implications for the unobserved errors in (25). In the case of v_{it}, for example, if the unknown output distance metafunction is a CD function and the researcher just happens to choose a CD approximating function, then

$$v_{it} = \left(\ln Q(q_{it}) - \sum_{n=1}^{N^*} \alpha_{nit} \ln q_{nit} \right) + \sum_{j=J+1}^{J^*} \rho_j \ln z_{jit} + \sum_{m=M+1}^{M^*} \beta_m \ln x_{mit}. \tag{26}$$

The term in parentheses can be viewed as a measurement error, and the remaining two terms are omitted variable errors. In the case of u_{it}, if firms maximise revenue and output markets are not perfectly competitive, then, irrespective of the properties of the metatechnology, $u_{it} = U^t(x_{it}, d_{it}, z_{it}) \geq 0$ (i.e., the inefficiency effect is a time-varying function of inputs, demand shifters and environmental variables).

In practice, there is generally no knowing for certain whether any assumptions concerning technologies, markets or behaviour are true. The usual way forward is to make such assumptions implicitly, by way of assumptions concerning the variables in (25). Standard statistical tests are then used to assess whether these assumptions are true. The most common assumptions are:

E1 $\varepsilon_{it} \equiv v_{it} - u_{it}$ is a random variable with $E(\varepsilon_{it}) = -\mu \leq 0$,

E2 $Cov(\varepsilon_{it}, x_{mit}) = Cov(\varepsilon_{it}, z_{jit}) = 0$ for all $m = 1, \ldots, M$ and $j = 1, \ldots, J$,

E3 $Var(\varepsilon_{it}) = \sigma^2$,

E4 $Cov(\varepsilon_{it}, \varepsilon_{ks}) = 0$ if $i \neq k$ or $t \neq s$,

E5 v_{it} is distributed as a normal random variable,

E6 u_{it} is distributed as a half-normal random variable,

E7 $u_{it} = \delta_0 + \sum_{j=1}^{J} \delta_j z_{jit} + \omega_{it} \geq 0$, and

E8 $u_{it} = \exp[-\eta(t - T)]u_i \geq 0$ where $u_i \geq 0$ and T is the last period in the sample.

The key assumptions are E1 and E2. Assumption E1 says the composite errors are random variables with the same mean. E2 says they are uncorrelated with the observed inputs and environmental variables. If these two assumptions are true then the OLS estimator is an unbiased and consistent estimator of the slope parameters in (25). It is common to implicitly assess the validity of E1 using standard model specification tests (e.g., testing for fixed time and/or firm effects). Assumption E2 can be tested using a Hausman test. If E2 is not satisfied, then consistent estimates of the slope parameters in (25) can be obtained using, for example, two-stage least squares (2SLS). The 2SLS estimator is a member of the family of instrumental variables estimators. In practice, knowledge of markets and firm behaviour can be used to guide the selection of instruments. For example, if input markets are perfectly competitive and firms minimise cost, then input prices are likely to be strong instruments (perfect competition means input prices are not affected by the actions of individual firms, and cost minimisation means they tend to be highly correlated with input quantities).

Assumptions E3 to E6 are important because they have implications for the efficiency of different econometric estimators. For example, if assumptions E1 to E4 are all true, then the OLS estimator (of the slope parameters) is the most efficient estimator in the class of linear unbiased estimators. If assumptions E1 to E6 are all true, then consistent and asymptotically efficient estimates can be obtained using the method of maximum likelihood (ML).

Assumption E7 can be traced back to Battese and Coelli (1995). This assumption is always true, but it generally brings nothing to the party unless the researcher has first made a poor choice of functional form. To see this, suppose there is only one input, one output, one environmental variable, and the log-metafunction is $\ln D^t_O(x_{it}, q_{it}, z_{it}) = \ln q_{it} - \gamma_t - \rho \ln z_{it} - \delta z_{it} - \beta \ln x_{it}$. In this case, E7 can be written as $u_{it} = \delta z_{it} + \omega_{it} \geq 0$ where $\omega_{it} = \gamma_t + \rho \ln z_{it} + \beta \ln x_{it} - \ln q_{it}$. If all variables are observed and the researcher (correctly) chooses $f(z_{it}, x_{it}, t) = \gamma_t + \rho \ln z_{it} + \delta z_{it} + \beta \ln x_{it}$, then (25) takes the form

$$\ln q_{it} = \gamma_t + \rho \ln z_{it} + \delta z_{it} + \beta \ln x_{it} + v_{it} - u_{it} \tag{27}$$

where $v_{it} = 0$. In this case, the ML estimator is consistent with or without E7 (i.e., E7 contains no additional information). On the other hand, if the researcher chooses a CD approximating function then (25) takes the form

$$\ln q_{it} = \gamma_t + \rho \ln z_{it} + \beta \ln x_{it} + v_{it} - u_{it} \tag{28}$$

where $v_{it} = \delta z_{it}$. In this case, neither E1 nor E2 are true and, without E7, the ML estimator is inconsistent. It is easily shown that (28) and E7 are together equivalent to (27) (i.e., E7 contains additional information). In this case, if the researcher chooses a CD approximating function and E7 is maintained, then the ML estimator is consistent.

Finally, assumption E8 is an implausible assumption that can be traced back to Battese and Coelli (1992). In practice, $u_i \geq 0$ is generally interpreted as a time-invariant technical inefficiency effect. This is implausible because, given the definition of u_{it}, it implies there is no statistical noise (i.e., it implies that all variables involved in the production process are observed, there is no measurement error, and the functional form of the output distance metafunction is known). Even if u_i is interpreted more broadly, as a combined measure of technical inefficiency and statistical noise, E8 cannot be true unless statistical noise is time-invariant. Thus, in practice, maintaining assumption E8 generally leads to biased and inconsistent estimates of OTE.

9. Empirical Illustration

This section uses publicly-available panel data to estimate agricultural efficiency and productivity change for $I = 11$ states in the northeastern U.S. for the $T = 30$ years from 1960 to 1989.[18] The main purpose is to illustrate the way in which non-sample information concerning data, technologies, markets and behaviour can be used to guide the estimation of efficiency and productivity change. Readers who are particularly interested in U.S. agricultural productivity should see Ball et al. (2004) and/or O'Donnell (2012*d*).

9.1. Data

The Economic Research Service (ERS) of the U.S. Department of Agriculture (USDA) is one of the most reliable public sources of state-level data on production and input use in U.S. agriculture. This section uses ERS data on $N = 3$ aggregate outputs (livestock, crops, other outputs), $M = 4$ aggregate inputs (capital, labour, land, materials) and $J = 1$ environmental variable (temperature) to compute and decompose a GY TFP index. Like most statistical agencies, the ERS uses quantity index formulas that do not satisfy a circularity axiom.[19] This is a common source of statistical noise: failure to satisfy the circularity axiom means, for example, that two states could produce exactly the same outputs and yet the output quantity index would say that their outputs were different.[20] Omitted variables (e.g., soil quality, rainfall) and measurement errors (e.g., aggregation from the farm-level to the state-level) are additional sources of statistical noise in the ERS accounts. This has obvious implications for the choice of estimator: recall that "if the data contain noise, ... there seems little choice but to rely on SFA" (Simar and Wilson, 2000, p. 77, fn. 14).

[18]The states are CT, DE, MA, MD, ME, NH, NH, NY, PA, RI and VT. The data are a subset of the data used by O'Donnell (2012*d*). Only a subset of those data are used here because, at the time of writing, the full dataset was not publicly available. Besides, this section is illustrative.

[19]The quantity indexes are implicit indexes computed by dividing values by EKS price indexes. More details concerning the construction of the variables can be found in Ball et al. (1997, 2004).

[20]Other indexes that fail the circularity axiom include chained Törnqvist and EKS indexes. These indexes are used by most statistical agencies, not just the ERS. For more details on index number axioms, see O'Donnell (2012*b*).

9.2. Technologies

In the case of U.S. agriculture, there is no reason to suppose that the period-t agricultural production metatechnology cannot be represented by an output distance metafunction. In that case, the relationship between the variables in the ERS dataset can be written in the form:

$$\ln Q_{it} = \gamma_0 + \gamma_1 t + \rho \ln z_{it} + \sum_{m=1}^{M} \beta_m \ln x_{mit} + v_{it} - u_{it} \tag{29}$$

where Q_{it} is (proportional to) the GY quantity index in equation (18). If statistical noise is random, then equation (29) is a CD stochastic production frontier. The noise component in (29) accounts for the possibility that there is measurement error in the construction of Q_{it}, as well as the possibility that the output distance metafunction is not, in fact, a CD function. There is no knowing whether these particular errors are random or deterministic (or zero). However, the noise component also subsumes omitted variable errors, and we can be confident that at least one omitted variable is random (e.g., rainfall). We can also be reasonably confident that agricultural inputs and outputs are strongly disposable. Agricultural nonpoint source pollution, for example, is evidence that inputs are strongly disposable. Thus, we can expect $\beta_1, \ldots, \beta_4 \geq 0$.

9.3. Markets and Behaviour

Four generally-accepted facts about U.S. agriculture are that i) total land area is fixed, ii) farmers are price-takers in input and output markets, iii) farmers make input decisions before output prices are known, and iii) they choose variable inputs in order to maximise expected profits. If land area is fixed, then, for all intents and purposes, it can be treated as a characteristic of the production environment.[21] If input prices and characteristics of the production environment are known at the time input decisions are made, then the period-t optimisation problem of farmer i is $V = \max_{x \geq 0, q \geq 0}\{p_{it}^{e\prime} q - w_{it}^{\prime} x : D_O^t(x, q, z_{it}) \leq 1\}$ where $p_{it}^{e} \geq 0$ is a vector of expected output prices. The output and input vectors that solve this problem are $q_{it}^{*} = q^{t}(p_{it}^{e}, w_{it}, z_{it})$ and $x_{it}^{*} = x^{t}(p_{it}^{e}, w_{it}, z_{it})$.

[21] In the case of production frontiers, fixed inputs and measures of quality are mathematically indistinguishable from characteristics of the production environment. To see this, simply redefine the production possibilities set as $T(z,g) = \{(x,q) :$ variable inputs x and fixed inputs z and technology g can produce outputs $q\}$ or $T(z,g) = \{(x,q) :$ technology g and inputs x of quality z can produce outputs $q\}$.

If the farmer forms naive expectations (i.e., if $p^e_{it} = p_{i,t-1}$) and chooses $q_{it} = q^*_{it}$ and $x_{it} = x^*_{it}$ then, among other things, $x_{it} = x^t(p_{i,t-1}, w_{it}, z_{it})$. In the present context, this means we can expect inputs of capital, labour and materials to be functions of lagged output prices and current period input prices, all of which are exogenous.

9.4. Parameters and Technical Efficiency

If assumptions E1 and E2 in Section 8 are true, then the OLS estimator is a consistent estimator of the slope parameters in (29). Moreover, a consistent estimator of OTE is a corrected least squares (CLS) estimator that can be traced back to Winston (1957, p.283).[22] The OLS parameter estimates and the average of the CLS OTE estimates are reported in the first column of Table 2. The parameter estimates are plausible: all of the coefficients have the expected signs and most are statistically significant at the 5% level; the coefficient of the time trend indicates that the sector has experienced technical progress at a rate of 1.1% per annum, which is a slightly lower rate than the estimated 1.8% per annum reported by Ray (1982); the estimated elasticity of scale is slightly less than one, which implies that a solution to the expected profit maximisation problem exists; and the coefficient of determination (not reported in Table 2) is 0.994, indicating that the CD function provides an excellent approximation to the unknown output distance metafunction. However, a Hausman test leads to the conclusion that the log-capital, log-labour and log-materials variables are endogenous (the p-value was less than 0.0001). Following the discussion in Section 9.3, this test was conducted using the lagged aggregate output price and the four input prices as instruments. Eleven observations (i.e., all the observations for 1960) were lost due to lagging. The practical significance of the Hausman test result is that the OLS estimator is almost certainly biased and inconsistent, meaning the parameter estimates and the associated average OTE score reported in the first column of Table 2 are unreliable.

[Table 2 near here]

The second column of Table 2 reports 2SLS estimates. Again, the lagged aggregate output price and four input prices were used as instruments. A test of over-identifying

[22]The CLS estimator of OTE_{it} is $\exp(\hat{e}_{it} - \hat{\mu})$ where $\hat{e}_{it}$ is the OLS residual and $\hat{\mu} = \max\{\hat{e}_{11}, \ldots, \hat{e}_{IT}\}$. Thus, the CLS estimator of OTE_{it}/OTE_{ks} does not depend on $\hat{\mu}$.

restrictions was used to confirm that these instruments are valid. The parameter estimates are generally plausible: one coefficient does not have the expected sign but it is statistically insignificant; the coefficient of the time trend indicates that the sector experienced technical progress at a rate of 1.2% per annum, which is only marginally higher than the OLS estimate; the estimated elasticity of scale is 0.976 and significantly different from one at levels of significance greater than 8.1% (the p-value is 0.081); and the coefficient of determination is 0.957, indicating that the model still explains most of the variation in observed log-outputs. One reason we might not use these 2SLS estimates to draw inferences concerning efficiency and productivity is that we might firmly believe that inputs are strongly disposable. In that case, we would want the estimated CD function to have the same properties as the metafunction it approximates. That is, we would want the estimates of $\beta_1, \ldots, \beta_4$ to be nonnegative. The third column of Table 2 reports 2SLS estimates that satisfy these nonnegativity restrictions. The restrictions were imposed using a simple accept-reject algorithm proposed by Geweke (1986) (this particular algorithm was used because it is simple to implement in the SHAZAM software package). The so-called restricted two-stage least squares (R2SLS) estimates are theoretically plausible, by construction. The R2SLS standard errors are much smaller than the 2SLS standard errors, indicating that the information content in the SD constraints is high. The CLS estimator was again used to estimate OTE. The R2SLS estimates of OTE are on average higher than the 2SLS estimates, but still lower than averages reported elsewhere in the literature [e.g., 0.88 by Morrison Paul and Nehring (2005)].[23]

The fourth column in Table 2 reports two-stage ML (2SML) estimates obtained under the assumption that the noise and inefficiency effects are distributed as normal and half-normal random variables respectively (i.e., assumptions E5 and E6). The first stage was identical to the first stage in 2SLS: each endogenous explanatory variable was regressed on all instrumental and exogenous explanatory variables, and the predictions from these regressions were saved (in effect, OLS was used to estimate three input demand equations). In the second stage, the endogenous variables in (29) were replaced by their predicted values, and the resulting equation was estimated by ML. Thus, the

[23]Morrison Paul and Nehring (2005) used an estimator proposed by Battese and Coelli (1988) (hereafter, BC88). The BC88 estimator usually yields much higher estimates of OTE than the CLS estimator used here.

only difference between the 2SML estimator and the 2SLS estimator is that in the second stage the model is estimated by ML instead of OLS. The 2SML parameter estimates are generally very different from the 2SLS estimates, indicating that the information content in E5 and E6 is high. However, the 2SML estimates of OTE are, on average, only slightly lower than the least squares estimates. These 2SML efficiency estimates were obtained using an estimator proposed by Battese and Coelli (1988) (hereafter BC88) (this estimator is consistent with E5 and E6 and is hard-wired into the FRONTIER software package).[24]

Finally, the columns labelled BC95 and BC92 report 2SML estimates obtained under assumptions E7 and E8. The BC95 estimates are similar to the 2SML estimates, indicating that the information content in E7 is small. The BC92 estimates are quite different to all other estimates, reflecting the fact that E8 cannot be true. The BC95 and BC92 estimates of OTE were obtained using estimators proposed by Battese and Coelli (1995, 1992) (these estimators are also hard-wired into the FRONTIER package).[25]

9.5. TFP Change and Efficiency Change

Irrespective of characteristics of technologies, markets and behaviour, the GY TFP index (18) is well defined. If the researcher writes the metafunction in the form of (29) then the GY TFP index can be decomposed as:

$$TFPI_{ksit}^{GY} = \left[\frac{\exp(\gamma_1 t)}{\exp(\gamma_1 s)}\right]\left[\frac{\exp(\rho \ln z_{it})}{\exp(\rho \ln z_{ks})}\right]\left[\frac{\exp(-u_{it})}{\exp(-u_{ks})}\right]\prod_{m=1}^{M}\left(\frac{x_{mks}}{x_{mit}}\right)^{\tau_m}\left[\frac{\exp(v_{it})}{\exp(v_{ks})}\right] \quad (30)$$

where $\tau_m \equiv \beta_m - \bar{s}_m$. This is an empirical version of equation (19). The first term measures technical change, the second term measures EE change, the third term measures OTE change, and, even though it is not a function of outputs, the third term (not in brackets) measures OSME change. The last term measures the change in statistical noise. Thus, equation (30) says that the GY TFP index can be decomposed as dTFP $= d$T $\times$ dEE $\times$ dOTE $\times$ dOSME $\times$ dNOISE. Evaluating these components is straightforward: in this paper, the dT, dEE and dOSME components were evaluated using the

[24]If the 2SML parameter estimates are used to compute CLS estimates of OTE then the average is an implausibly low 0.469.

[25]If the BC95 and BC92 parameter estimates are used to compute CLS estimates of OTE then the averages are 0.427 and 0.553 respectively.

parameter estimates reported in Table 2; dOTE was evaluated by computing changes in estimates of OTE; and dNOISE was computed as a residual. Importantly, it is easily shown that if the CLS estimator is used to estimate OTE then $d\text{NOISE} = 1$. For purposes of comparison, selected components were also estimated using DEA: in this paper, DEA estimates of dT were obtained by computing changes in the maximum TFP in each time period; dOTE was obtained by solving LP (22);[26] dNOISE was set to one; and $d\text{EE} \times d\text{OSME}$ was computed as a residual.

Table 3 compares (estimated) levels of TFP and efficiency in Connecticut (CT), Delaware (DE) and Maryland (MD) in 1989 with corresponding levels in CT in 1961. The interpretation of these estimates is straightforward. For example, the first entry in the OLS column indicates that CT farmers were 47.5% more productive in 1989 than they had been in 1961 ($d\text{TFP} = 1.475$).[27] The next five entries in this column indicate that this was due to a combination of significant technical progress ($d\text{T} = 1.357$), a 1.3% improvement in characteristics of the production environment ($d\text{EE} = 1.013$), a 7.6% increase in scale-mix efficiency ($d\text{OSME} = 1.076$), and a 0.2% fall in technical efficiency ($d\text{OTE} = 0.998$). The noise component vanishes by construction. The different components in (30) are transitive, so the estimates in Table 3 can be used to make meaningful comparisons of TFP and efficiency across both states and time. For example, the entries in the OLS column indicate that DE farmers were 39% more productive than CT farmers in 1989 ($d\text{TFP} = 2.050/1.475 = 1.390$), mainly because they were 27.7% more scale-mix efficient ($d\text{OSME} = 1.374/1.076 = 1.277$).

[Table 3 near here]

Finally, the estimates reported in Table 3 reveal that different estimators lead to qualitatively different estimates of the efficiency change components productivity change. To get a clearer picture of these differences, Figure 1 presents measures of TFP and efficiency change in MD from 1961 to 1989 (cf. CT in 1961). Observe that the DEA and OLS estimators paint plausible but quite different pictures of the drivers of TFP change: the DEA estimator suggests that the only source of TFP growth was technical

[26]For computational simplicity, the output distance metafunction was assumed to be nonincreasing in temperature. This is equivalent to treating temperature as an input.

[27]Note that there are no unknown parameters in (18), so estimates of dTFP do not depend on the choice of econometric estimator. However, the components of TFP change are sensitive to this choice.

progress, while the OLS estimator suggest that TFP growth was due to a combination of technical progress and improvements in EE, OSME and OTE. The 2SLS and R2SLS results are similar to the OLS results, indicating that the endogeneity problem is relatively unimportant when it comes to estimating the components of productivity change (in this empirical application). The 2SML and BC95 estimators paint plausible but inaccurate (i.e., noisy) pictures of TFP change. This suggests that the normality and half-normality assumptions underpinning these estimators may not be true. Finally, the BC92 estimator paints a noisy picture in which OTE is steadily, and implausibly, falling. This reflects the fact that the assumptions underpinning this model cannot be true.

[Figure 1 near here]

10. Conclusion

Measuring TFP change is an important part of economic decision-making. Different decision-makers choose different measures of TFP change (i.e., different index formulas) depending on what they value. For example, most national statistical agencies construct TFP indexes by assigning zero weights to goods and services that do not have a market price, and in so doing they place a value of zero on, for example, unpriced household production (e.g., cooking, cleaning, childcare) and environmental bads (i.e., pollution). This paper considers a new TFP index formula that does not involve the use of market prices: it only involves vectors of input quantities, output quantities and environmental variables. The advantage of this index is that it can be computed when prices are either unavailable (e.g., because they have not been collected) or negative (e.g., bad outputs). The downside is that it cannot be computed without estimating a functional representation of the metatechnology (e.g., an output distance metafunction). One of the contributions of the paper is to show how non-sample information concerning technologies, markets and firm behaviour can be used to inform the estimation process. For example, the paper points out that if inputs and/or outputs are strongly disposable, then the output distance metafunction cannot be a translog function. The practical implication is that if inputs and/or outputs are strongly disposable, then a translog production frontier model must contain an error term representing statistical noise. The paper also shows that if the output distance metafunction is a Cobb-Douglas function, technical change is Hicks neutral, markets are perfectly competitive, and firms maximise profit, then the new TFP index proposed in the paper collapses to the GY TFP index defined by O'Donnell (2011). The GY TFP index is an index that uses sample average revenue and cost shares as weights, so it assigns zero weights to goods and services that do not have a market price.

Measuring (or estimating) TFP change is only a small part of productivity analysis. A much larger and more difficult part is estimating the drivers of TFP change. Another of the contributions of the paper is to show how the GY TFP index can be decomposed into measures of technical change, environmental efficiency change, output-oriented technical efficiency change, and output-oriented scale-mix efficiency change. The ability to identify these components has important implications for public policy-making: as O'Donnell (2012*d*, p.873) puts it, if we cannot identify the drivers of TFP change, then "policy-makers cannot properly assess whether the payoffs from improving the rate of technical progress (e.g., through increased R&D expenditure) are more or less likely to outweigh the payoffs from improving levels of either technical efficiency (e.g., through education and training programs) or scale and mix efficiency (e.g., by using taxes and subsidies to change relative prices)." This paper shows how non-sample information concerning technologies, markets and firm behaviour can influence the accuracy with which these components are estimated. In an empirical application to U.S. farm data, least squares estimators painted the most plausible and accurate picture of the drivers of TFP change.

Finally, the theory and empirical results presented in this paper suggest that the DEA and BC92 estimators are biased and inconsistent estimators of OTE. This does not necessarily mean they are biased and inconsistent estimators of OTE change. Further investigation of this issue seems worthwhile.

Appendix

Proposition 1. *If a metatechnology is IHON then* $D_O^t(x,q,z) = D_O^{\bar{s}}(x,q,z)/a(x,z,\bar{s},t)$.

Proof. Let $\kappa = a(x,z,\bar{s},t)$. If IHON then $D_O^t(x,q,z) = \inf\{\delta > 0 : q/\delta \in P^t(x,z)\} = \inf\{\delta > 0 : q/\delta \in \kappa P^{\bar{s}}(x,z)\} = \inf\{\delta > 0 : q/(\delta\kappa) \in P^{\bar{s}}(x,z)\} = \inf\{(\delta\kappa)/\kappa > 0 : q/(\delta\kappa) \in P^{\bar{s}}(x,z)\} = (1/\kappa)\inf\{(\delta\kappa) > 0 : q/(\delta\kappa) \in P^{\bar{s}}(x,z)\}$
$= D_O^{\bar{s}}(x,q,z)/\kappa = D_O^{\bar{s}}(x,q,z)/a(x,z,\bar{s},t)$ □

Proposition 2. *If a metatechnology is OH then* $D_O^t(x,q,z) = D_O^t(\bar{x},q,\bar{z})/f(x,z,\bar{x},\bar{z},t)$.

Proof. Let $\kappa = f(x,z,\bar{x},\bar{z},t)$. If OH then $D_O^t(x,q,z) = \inf\{\delta > 0 : q/\delta \in P^t(x,z)\} = \inf\{\delta > 0 : q/\delta \in \kappa P^t(\bar{x},\bar{z})\} = \inf\{\delta > 0 : q/(\delta\kappa) \in P^t(\bar{x},\bar{z})\} = \inf\{(\delta\kappa)/\kappa > 0 : q/(\delta\kappa) \in P^t(\bar{x},\bar{z})\} = (1/\kappa)\inf\{(\delta\kappa) > 0 : q/(\delta\kappa) \in P^t(\bar{x},\bar{z})\}$
$= D_O^t(\bar{x},q,\bar{z})/\kappa = D_O^t(\bar{x},q,\bar{z})/f(x,z,\bar{x},\bar{z},t)$ □

Proposition 3. *If a metatechnology is IHIN then* $D_I^t(x,q,z) = D_I^{\bar{s}}(x,q,z)/b(q,z,\bar{s},t)$.

Proof. Let $\kappa = b(q,z,\bar{s},t)$. If IHIN then $D_I^t(x,q,z) = \sup\{\rho > 0 : x/\rho \in L^t(q,z)\} = \sup\{\rho > 0 : x/\rho \in \kappa L^{\bar{s}}(q,z)\} = \sup\{\rho > 0 : x/(\kappa\rho) \in L^{\bar{s}}(q,z)\} = \sup\{(\kappa\rho)/\kappa > 0 : x/(\kappa\rho) \in L^{\bar{s}}(q,z)\} = (1/\kappa)\sup\{(\kappa\rho) > 0 : x/(\kappa\rho) \in L^{\bar{s}}(q,z)\} = D_I^{\bar{s}}(x,q,z)/\kappa$
$= D_I^{\bar{s}}(x,q,z)/b(q,z,\bar{s},t)$ □

Proposition 4. *If a metatechnology is IH then* $D_I^t(x,q,z) = D_I^t(x,\bar{q},\bar{z})/h(q,z,\bar{q},\bar{z},t)$.

Proof. Let $\kappa = h(q,z,\bar{q},\bar{z},t)$. If IH then $D_I^t(x,q,z) = \sup\{\rho > 0 : x/\rho \in L^t(q,z)\} = \sup\{\rho > 0 : x/\rho \in \kappa L^t(\bar{q},\bar{z})\} = \sup\{\rho > 0 : x/(\kappa\rho) \in L^t(\bar{q},\bar{z})\} = \sup\{(\kappa\rho)/\kappa > 0 : x/(\kappa\rho) \in L^t(\bar{q},\bar{z})\} = (1/\kappa)\sup\{(\kappa\rho) > 0 : x/(\kappa\rho) \in L^t(\bar{q},\bar{z})\} = D_I^t(x,\bar{q},\bar{z})/\kappa$
$= D_I^t(x,\bar{q},\bar{z})/h(q,z,\bar{q},\bar{z},t)$ □

Proposition 5. *If a metatechnology is OH and technical change is IHON then* $D_O^t(x,q,z) = Q(q)/F^t(x,z)$ *where* $Q(q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})$ *and* $F^t(x,z) = a(\bar{x},\bar{z},\bar{s},t)f(x,z,\bar{x},\bar{z},t)$.

Proof. IHON $\Rightarrow D_O^t(x,q,z) = D_O^{\bar{s}}(x,q,z)/a(x,z,\bar{s},t) \Rightarrow D_O^t(\bar{x},q,\bar{z}) = D_O^{\bar{s}}(\bar{x},q,\bar{z})/a(\bar{x},\bar{z},\bar{s},t)$. OH $\Rightarrow D_O^t(x,q,z) = D_O^t(\bar{x},q,\bar{z})/f(x,z,\bar{x},\bar{z},t)$. Thus, IHON and OH $\Rightarrow D_O^t(x,q,z) = D_O^{\bar{s}}(\bar{x},q,\bar{z})/[a(\bar{x},\bar{z},\bar{s},t)f(x,z,\bar{x},\bar{z},t)] = Q(q)/F^t(x,z)$ □

Proposition 6. *If a metatechnology is IH and technical change is IHIN then* $D_I^t(x,q,z) = X(x)/H^t(q,z)$ *where* $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$ *and* $H^t(q,z) = b(\bar{q},\bar{z},\bar{s},t)h(q,z,\bar{q},\bar{z},t)$.

Proof. IHIN $\Rightarrow D_I^t(x,q,z) = D_I^{\bar{s}}(x,q,z)/b(q,z,\bar{s},t) \Rightarrow D_I^t(x,\bar{q},\bar{z}) = D_I^{\bar{s}}(x,\bar{q},\bar{z})/b(\bar{q},\bar{z},\bar{s},t)$. IH $\Rightarrow D_I^t(x,q,z) = D_I^t(x,\bar{q},\bar{z})/h(q,z,\bar{q},\bar{z},t)$. Thus, IHIN and IH $\Rightarrow D_I^t(x,q,z) = D_I^{\bar{s}}(x,\bar{q},\bar{z})/[b(\bar{q},\bar{z},\bar{s},t)h(q,z,\bar{q},\bar{z},t)] = X(x)/H^t(q,z)$ □

Proposition 7. *If technical change is HN and the metatechnology is H and HDr then* $D_O^t(x,q,z) = Q(q)/[A^t(z)X(x)^r]$ *where* $Q(q) = D_O^{\bar{s}}(\bar{x},q,\bar{z})$, $X(x) = D_I^{\bar{s}}(x,\bar{q},\bar{z})$ *and* $A^t(z) = D_O^{\bar{s}}(\bar{x},\bar{q},\bar{z})^2/D_O^t(\bar{x},\bar{q},z)$.

Proof. OH and IHON $\Rightarrow D_O^t(x,q,z) = Q(q)/F^t(x,z)$ (A). IH and IHIN $\Rightarrow D_I^t(x,q,z) = X(x)/H^t(q,z)$. HD$r \Rightarrow D_O^t(x,q,z) = D_I^t(x,q,z)^{-r}$. Thus, IH, IHIN and HD$r \Rightarrow D_O^t(x,q,z) = H^t(q,z)^r/X(x)^r$ (B). (A) $\Rightarrow D_O^t(x,\bar{q},z) = Q(\bar{q})/F^t(x,z)$ (C). (B) $\Rightarrow D_O^t(x,\bar{q},z) = H^t(\bar{q},z)^r/X(x)^r$ (D). (B) $\Rightarrow D_O^t(\bar{x},\bar{q},z) = H^t(\bar{q},z)^r/X(\bar{x})^r \Rightarrow H^t(\bar{q},z)^r = D_O^t(\bar{x},\bar{q},z)X(\bar{x})^r$ (E). (C) and (D) $\Rightarrow Q(\bar{q})/F^t(x,z) = H^t(\bar{q},z)^r/X(x)^r \Rightarrow F^t(x,z) = X(x)^rQ(\bar{q})/H^t(\bar{q},z)^r$ (F). (E) and (F) $\Rightarrow F^t(x,z) = X(x)^rQ(\bar{q})/[D_O^t(\bar{x},\bar{q},z)X(\bar{x})^r]$ (G). If $A^t(z) = D_O^{\bar{s}}(\bar{x},\bar{q},\bar{z})^2/D_O^t(\bar{x},\bar{q},z)$ then HD$r \Rightarrow A^t(z) = D_O^{\bar{s}}(\bar{x},\bar{q},\bar{z})D_I^{\bar{s}}(\bar{x},\bar{q},\bar{z})^{-r}/D_O^t(\bar{x},\bar{q},z) = Q(\bar{q})X(\bar{x})^{-r}/D_O^t(\bar{x},\bar{q},z) = Q(\bar{q})/[D_O^t(\bar{x},\bar{q},z)X(\bar{x})^r]$ (H). (G) and (H) $\Rightarrow F^t(x,z) = X(x)^rA^t(z)$ (I). (A) and (I) $\Rightarrow D_O^t(x,q,z) = Q(q)/[A^t(z)X(x)^r]$ □

Table 1: Selected Assumptions Concerning Metasets and Metafunctions

	Assumption	Definition
A7	outputs strongly disposable	$(x,q) \in T^t(z)$ and $0 \leq q^1 \leq q \Rightarrow (x,q^1) \in T^t(z)$
A8	inputs strongly disposable	$(x,q) \in T^t(z)$ and $x^1 \geq x \Rightarrow (x^1,q) \in T^t(z)$
A9	homogeneous of degree r	$(x,q) \in T^t(z) \Leftrightarrow (\lambda x, \lambda^r q) \in T^t(z)$ for all $\lambda > 0$
A10	imp. Hicks output neutral	$P^t(x,z) = a(x,z,\bar{s},t)P^{\bar{s}}(x,z)$ for any $\bar{s}$
A11	output homothetic	$P^t(x,z) = f(x,z,\bar{x},\bar{z},t)P^t(\bar{x},\bar{z})$ for any $\bar{x}$, $\bar{z}$
A12	imp. Hicks input neutral	$L^t(q,z) = b(q,z,\bar{s},t)L^{\bar{s}}(q,z)$ for any $\bar{s}$
A13	input homothetic	$L^t(q,z) = h(q,z,\bar{q},\bar{z},t)L^t(\bar{q},\bar{z})$ for any $\bar{q}$, $\bar{z}$
D7	outputs strongly disposable	$\partial D_O^t(x,q,z)/\partial q_n \geq 0$ for all n
D8	inputs strongly disposable	$\partial D_O^t(x,q,z)/\partial x_m \leq 0$ for all m
D9	homogeneous of degree r	$D_O^t(\lambda x,q,z) = \lambda^{-r} D_O^t(x,q,z)$ for all $\lambda > 0$
D10	imp. Hicks output neutral	$D_O^t(x,q,z) = D_O^{\bar{s}}(x,q,z)/a(x,z,\bar{s},t)$ for any $\bar{s}$
D11	output homothetic	$D_O^t(x,q,z) = D_O^t(\bar{x},q,\bar{z})/f(x,z,\bar{x},\bar{z},t)$ for any $\bar{x}$, $\bar{z}$
D12	imp. Hicks input neutral	$D_I^t(x,q,z) = D_I^{\bar{s}}(x,q,z)/b(q,z,\bar{s},t)$ for any $\bar{s}$
D13	input homothetic	$D_I^t(x,q,z) = D_I^t(x,\bar{q},\bar{z})/h(q,z,\bar{q},\bar{z},t)$ for any $\bar{q}$, $\bar{z}$

Table 2: Parameter Estimates†

		OLS	2SLS	R2SLS	2SML	BC95	BC92
γ_0	intercept	$\underline{-1.861}$	2.211	0.553	12.361	12.637	1.743
		(0.355)	(3.464)	(0.889)	(8.715)	(7.789)	(4.234)
γ_1	time	$\underline{0.011}$	$\underline{0.012}$	$\underline{0.011}$	$\underline{0.010}$	$\underline{0.012}$	$\underline{0.022}$
		(0.001)	(0.002)	(0.001)	(0.004)	(0.004)	(0.002)
ρ	degree days	$\underline{0.224}$	-0.232	-0.047	-1.229	-1.243	0.151
		(0.044)	(0.364)	(0.100)	(0.912)	(0.818)	(0.410)
β_1	capital	$\underline{0.105}$	0.001	$\underline{0.045}$	0.261	0.276	-0.077
		(0.039)	(0.151)	(0.037)	(0.342)	(0.317)	(0.144)
β_2	labour	$\underline{0.216}$	-0.012	$\underline{0.072}$	-0.889	-1.027	0.279
		(0.028)	(0.288)	(0.054)	(0.721)	(0.643)	(0.286)
β_3	land	0.016	0.356	$\underline{0.198}$	1.310	$\underline{1.592}$	0.282
		(0.023)	(0.317)	(0.090)	(0.815)	(0.723)	(0.320)
β_4	materials	$\underline{0.659}$	$\underline{0.632}$	$\underline{0.670}$	0.249	0.104	0.314
		(0.017)	(0.172)	(0.057)	(0.427)	(0.383)	(0.183)
δ_0	u-intercept					-4.979	
						(5.212)	
δ_1	u_degree days					0.002	
						(0.002)	
η	u_time						$\underline{-0.010}$
							(0.003)
Elasticity of scale		0.996	0.976	0.984	0.931	0.945	0.798
Average OTE		0.770	0.765	0.800	0.723	0.788	0.601

†Standard errors are in parentheses; underlining indicates significance at the 5% level; R2SLS estimates of $\beta_1, \ldots, \beta_4$ are significant, by construction.

Table 3: TFP and Efficiency Indexes: CT, DE & MD in 1989 cf. CT in 1961

Index	DEA	OLS	2SLS	R2SLS	2SML	BC95	BC92
Connecticut							
dTFP	1.475	1.475	1.475	1.475	1.475	1.475	1.475
dT	1.726	1.357	1.395	1.355	1.340	1.392	1.839
dEE		1.013	0.987	0.997	0.933	0.932	1.009
dOSME		1.076	1.038	1.068	1.006	0.932	0.982
$d\text{EE} \times d\text{OSME}$	0.839	1.089	1.024	1.065	0.938	0.869	0.991
dOTE	1.019	0.998	1.033	1.022	0.923	0.980	0.792
dNOISE	1	1	1	1	1.271	1.245	1.023
Delaware							
dTFP	2.050	2.050	2.050	2.050	2.050	2.050	2.050
dT	1.726	1.357	1.395	1.355	1.340	1.392	1.839
dEE		1.081	0.923	0.984	0.654	0.651	1.053
dOSME		1.374	1.558	1.553	1.691	1.469	0.892
$d\text{EE} \times d\text{OSME}$	1.122	1.484	1.438	1.528	1.106	0.957	0.939
dOTE	1.059	1.017	1.022	0.990	1.218	1.176	1.164
dNOISE	1	1	1	1	1.135	1.309	1.020
Maryland							
dTFP	1.657	1.657	1.657	1.657	1.657	1.657	1.657
dT	1.726	1.357	1.395	1.355	1.340	1.392	1.839
ddEE		1.071	0.931	0.986	0.686	0.683	1.047
dOSME		1.188	1.402	1.342	2.334	2.442	0.791
$d\text{EE} \times d\text{OSME}$	0.907	1.273	1.306	1.323	1.600	1.667	0.829
dOTE	1.059	0.959	0.910	0.925	0.815	0.627	1.084
dNOISE	1	1	1	1	0.948	1.139	1.003

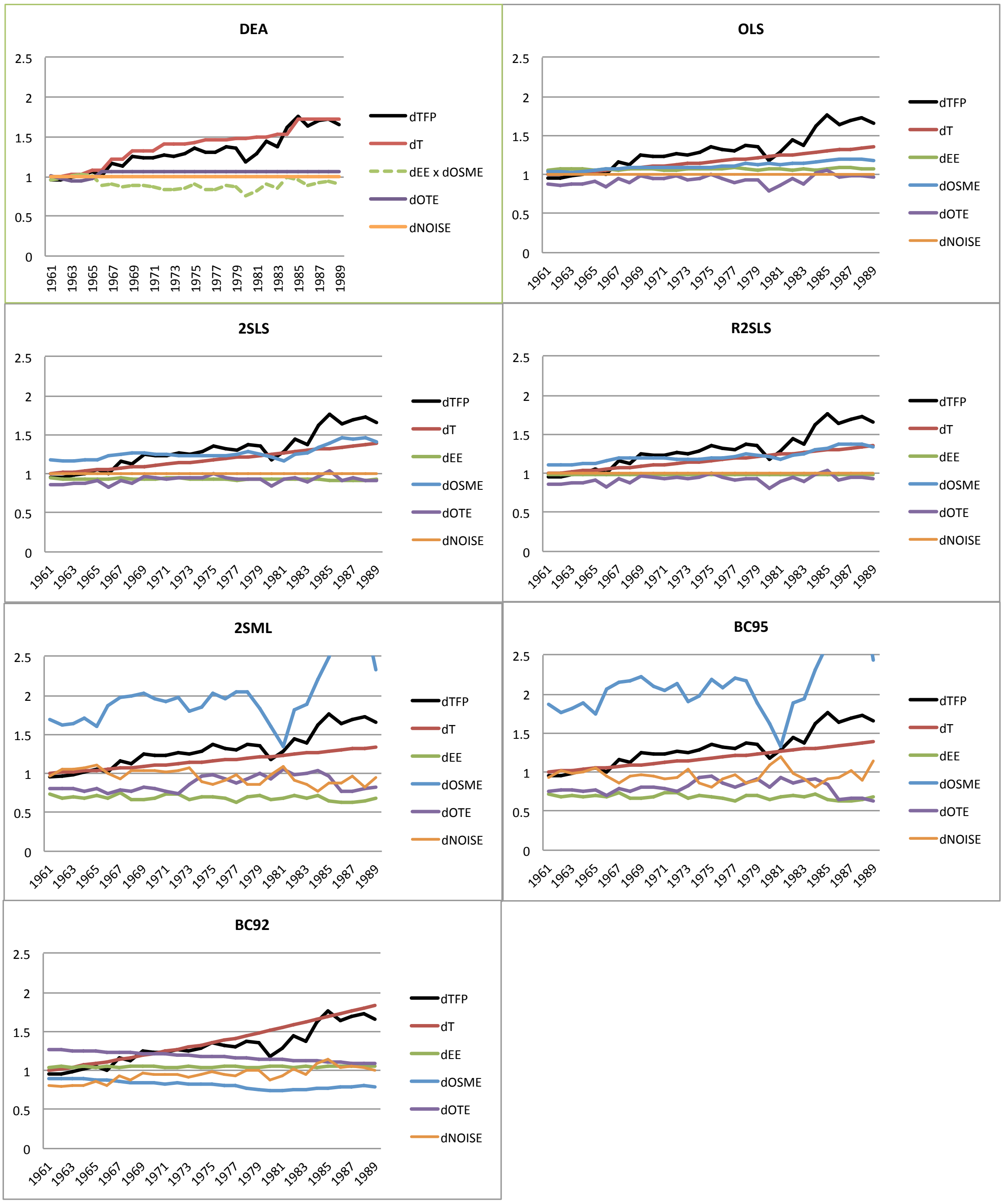

Figure 1: TFP and Efficiency Indexes: MD from 1961–1989 cf. CT in 1961

References

Afriat, S. (1972), 'Efficiency estimation of production functions', *International Economic Review* **13**(3), 568–598.

Aigner, D. and Chu, S. (1968), 'On estimating the industry production function', *The American Economic Review* **58**(4), 826–839.

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

Balk, B. (1998), *Industrial Price, Quantity, and Productivity Indices: The Micro-Economic Theory and an Application*, Kluwer Academic Publishers, Boston.

Balk, B. (2008), *Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference*, Cambridge University Press, New York.

Ball, V., Bureau, J.-C., Nehring, R. and Somwaru, A. (1997), 'Agricultural productivity revisited', *American Journal of Agricultural Economics* **79**(4), 1045–1063.

Ball, V., Hallahan, C. and Nehring, R. (2004), 'Convergence of productivity: An analysis of the catch-up hypothesis within a panel of states', *American Journal of Agricultural Economics* **86**(5), 1315–1321.

Banker, R., Charnes, A. and Cooper, W. (1984), 'Some models for estimating technical and scale inefficiencies in data envelopment analysis', *Management Science* **30**(9), 1078–1092.

Banker, R. and Morey, R. (1986), 'Efficiency analysis for exogenously fixed inputs and outputs', *Operations Research* **34**(4), 513–521.

Battese, G. and Coelli, T. (1988), 'Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data', *Journal of Econometrics* **38**, 387–399.

Battese, G. and Coelli, T. (1992), 'Frontier production functions, technical efficiency and panel data: With application to paddy farmers in india', *Journal of Productivity Analysis* **3**(1-2), 153–169.

Battese, G. and Coelli, T. (1995), ‘A model for technical inefficiency effects in a stochastic frontier production function for panel data’, *Empirical Economics* **20**, 325–332.

Chambers, R. and Quiggin, J. (2000), *Uncertainty, Production, Choice and Agency: The State-Contingent Approach*, Cambridge University Press, Cambridge, UK.

Charnes, A., Cooper, W. and Rhodes, E. (1978), ‘Measuring the efficiency of decision-making units’, *European Journal of Operational Research* **2**(6), 429–444. See also “Corrections” op.cit. 3(4):339.

Christensen, L., Jorgenson, D. and Lau, L. (1973), ‘Transcendental logarithmic production frontiers’, *The Review of Economics and Statistics* **55**(1), 28–45.

Cobb, C. and Douglas, P. (1928), ‘A theory of production’, *The American Economic Review* **18**(1), 139–165.

Debreu, G. (1951), ‘The coefficient of resource utilization’, *Econometrica* **19**(3), 273–292.

Färe, R. and Primont, D. (1995), *Multi-output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.

Farrell, M. (1957), ‘The measurement of productive efficiency’, *Journal of the Royal Statistical Society, Series A (General)* **120**(3), 253–290.

Geweke, J. (1986), ‘Exact inference in the inequality constrained normal linear regression model’, *Journal of Applied Econometrics* **1**, 127–141.

Heady, E. and Dillon, J. (1961), *Agricultural Production Functions*, Iowa State University Press, Ames, Iowa.

Jorgenson, D. and Griliches, Z. (1967), ‘The explanation of productivity change’, *The Review of Economic Studies* **34**(3), 249–283.

Kneip, A., Simar, L. and Wilson, P. (2011), ‘A computationally efficient, consistent bootstrap for inference with non-parametric DEA estimators’, *Computational Economics* **38**(4), 483–515.

Leibenstein, H. (1979), 'X-efficiency: From concept to theory', *Challenge* **22**(4), 13–22.

Meeusen, W. and van den Broeck, J. (1977), 'Efficiency estimation from Cobb-Douglas production functions with composed error', *International Economic Review* **18**(2), 435–444.

Morrison Paul, C. and Nehring, R. (2005), 'Product diversification, production systems, and economic performance in u.s. agricultural production', *Journal of Econometrics* **126**(2), 525–548.

Nadiri, M. (1970), 'Some approaches to the theory and measurement of total factor productivity: A survey', *Journal of Economic Literature* **8**(4), 1137–1177.

O'Donnell, C. (2012*a*), 'An aggregate quantity framework for measuring and decomposing productivity change', *Journal of Productivity Analysis* **38**(3), 255–272.

O'Donnell, C. (2012*b*), Alternative indexes for multiple comparisons of quantities and prices, Centre for Efficiency and Productivity Analysis Working Papers WP05/2012 (Version 21 May 2013), University of Queensland.

O'Donnell, C. (2012*c*), 'Econometric estimation of distance functions and associated measures of productivity and efficiency change', *Journal of Productivity Analysis* **DOI: 10.1007/s11123-012-0311-1**.

O'Donnell, C. (2012*d*), 'Nonparametric estimates of the components of productivity and profitability change in U.S. agriculture', *American Journal of Agricultural Economics* **94**(4), 873–890.

O'Donnell, C. J. (2011), The sources of productivity change in the manufacturing sectors of the U.S. economy, Centre for Efficiency and Productivity Analysis Working Papers WP07/2011, University of Queensland.

O'Donnell, C. and Nguyen, K. (2013), 'An econometric approach to estimating support prices and measures of productivity change in public hospitals', *Journal of Productivity Analysis* **40**(3), 323–335.

O'Donnell, C., Rao, D. and Battese, G. (2008), 'Metafrontier frameworks for the study of firm-level efficiencies and technology ratios', *Empirical Economics* **34**(2), 231–255.

Ray, S. (1982), 'A translog cost function analysis of u.s. agriculture, 1939-77', *American Journal of Agricultural Economics* **64**(3), 490–498.

Shephard, R. (1970), *The Theory of Cost and Production Functions*, Princeton University Press, Princeton.

Simar, L. and Wilson, P. (2000), 'Statistical inference in nonparametric frontier models: The state of the art', *Journal of Productivity Analysis* **13**(1), 49–78.

Solow, R. (1957), 'Technical change and the aggregate production function', *Review of Economics and Statistics* **39**(3), 312–320.

Winston, C. (1957), 'Discussion on Mr. Farrell's paper', *Journal of the Royal Statistical Society, Series A (General)* **120**(3), 282–84.