Productivity Growth in Industrialized Countries

Léopold Simar

and

Paul W. Wilson*

February 1998

*Simar: Institut de Statistique and CORE, Université Catholique de Louvain, Voie du Roman Pays 20, Louvain-la-Neuve, Belgium. Wilson: Department of Economics, University of Texas, Austin, Texas 78712 USA. We have benefited from conversations with James F. Burgess, Jr. and David C. Wheelock. We are grateful for financial support from the Belgian Government ("Projet d'Actions de Recherche Concertées, PARC no. 93/98–164), and from the Management Science Group, US Department of Veterans Affairs. We alone, of course, are responsible for any remaining errors in the current text.

Several recent papers in this *Review* examined important questions regarding productivity growth and its sources in industrialized countries. Among other findings, Färe, Grosskopf, Norris, and Zhang (1994; hereafter FGNZ) reported an estimated average annual growth rate of 0.85 percent for productivity in the US over 1979–1988. By contrast, Ray and Desli (1997; herafter RD) reported a corresponding estimate of −5.56 percent per annum over 1979–90. The difference in these estimates of the annual productity growth rate—6.41 percent—is striking. Moreover, it is implausible that US productivity declined by 5.56 percent per annum during 1979–90, particularly since US gross domestic product increased by 2.3 percent per year over this period while the rate of growth in the ratio of gross domestic product to labor exceeded the growth rate of the capital/labor ratio in the US over this period. In their reply to RD, Färe, Grosskopf, and Norris (1997; herafter FGN) were silent on this huge discrepancy.

We examine two sets of issues raised by these papers, and reassess what can be learned about productivity, efficiency, and technology from the data used by both papers. The first set of issues are primarily economic in nature. FGNZ stated in their abstract that

> "This paper analyzes productivity growth in 17 OECD countries over the period 1979–1988. A nonparametric programming method (activity analysis) is used to compute Malmquist productivity indexes. These are decomposed into two component measures, namely, technical change and efficiency change."

The measure of efficiency change was subsequently decomposed into measures of "pure efficiency change" and change in scale efficiency. RD were rather critical of the approach taken by FGNZ, and stated that the decomposition of productivity change into changes in efficiency, scale, and technology used by FGNZ is "internally inconsistent." RD offered an alternative decomposition of the Malmquist index in which the FGNZ measure of pure efficiency change appears, but in which both the scale efficiency change and technical change measures differ. They claimed (in the last sentence of their introduction) that use of their assumptions "leads one to *significantly* different conclusions" (italics added by present authors) than those obtained by FGNZ. But, in RD's alternative decomposition, the component which is supposed to measure changes in returns to scale confounds the different effects of movement of production units in input/output space and changes in the shape of the technology over time. FGN recognized that RD's decomposition incorrectly measures changes in scale efficiency, but not that RD's measure confounds two different effects. We offer an alternative decomposition below which avoids this problem.

The second set of issues we examine concern estimation and inference. We find that many of the values RD reported in Tables 1–2 of their paper, including their estimate of average annual productivity growth for the US, are incorrect. Aside from this problem, the various "measures" discussed by FGNZ and RD were computed from data, and as such are *estimates* of *true,* but unobserved, quantities. Neither FGNZ nor RD defined a statistical model, although one is implied by their theoretical discussions as well as their empirical applications. Neither paper considered the properties of the *nonparametric, statistical estimators* they used; neither distinguished between unobserved true quantities and estimates of these quantities. Compounding the problem, both papers drew numerous inferences from their empirical results, although they had no statistical basis for doing so. A final point, related to the previous one, is that both papers used nonparametric estimators for which only the *asymptotic* rates of convergence are known; the estimators' small-sample properties remain unknown, and both papers estimated various quantities from observations on only 17 countries. It is quite possible that the estimates of productivity change, efficiency change, *etc.* reported in both papers have large variances. This casts doubt on the claim by RD that their results are *significantly* different from those of FGNZ, particularly since RD did not provide statistical tests of significance and since their estimates were, in many cases, numerically close to those obtained by FGNZ We provide a statistical model suggested by the original framework of FGNZ which allows us to estimate confidence intervals and formally test many of the claims made by both papers.

FGNZ note that their linear programming methods are similar to the nonparametric methods used in Chavas and Cox (1990), which are also based on linear programming methods. Indeed, much of the discussion in Chavas and Cox seems to assume that the technology is known with certainty, when in fact it must be estimated. Thus, many of the above comments also apply to the paper by Chavas and Cox, as well as the papers by FGNZ and by RD. Similar comments also apply to the papers cited by FGN offering alternative decompositions of productivity indices.

One could argue that FGNZ and RD merely wanted to make relative comparisons among the 17 OECD countries they included in their analyses, and that thus they had no need to consider econometric issues. Neither paper made such a claim. Moreover, this view is refuted by statements in both papers. For example, FGNZ wisely cautioned (page 81) that

> "Our results here should be interpreted with care. Our sample of countries is arbitrary—they are the countries for which we were able to collect consistent data over this period."

Indeed, it is quite likely that any comparison of say, the United States and Japan, would have been different if a different sample had been used. As noted earlier, RD stated that their results were *significantly* different from those of FGNZ, which would usually imply a statistical result. Moreover, in discussing their results, RD made a series of specific statements about rates of technical change experienced by several countries—*i.e.,* they drew inferences from their results. For example, they stated (first paragraph of their section II) that

> "The overall average of the [constant-returns-to-scale]-based technical change index was 0.99303. This implies a very slight rate of technical regress. On the other hand, the average of [variable-returns-to-scale]-based technical change was 1.00367 showing technical progress at the rate of 0.37% per year."

Aside from the question of whether returns to scale are constant or variable, the question of whether either of these numbers are significantly different from unity in a statistical sense is of paramount importance. In other words, are the differences from unity due only to the fact that a small, arbitrary sample was used, or do they imply real economic phenomena? One might also wonder whether the differences from unity are significant in an economic sense; apparently, RD believed they were. However, the question of *economic* significance can only be relevant if the differences are found to be significant in a *statistical* sense.

I. Economic Issues

Many of the inaccuracies in FGNZ, RD, and FGN may be attributed to their confusion between unknown quantities and estimates of these quantities. We introduce some new notation in order to precisely describe what the previous authors did (or did not do) and to clarify both the economic and the econometric issues.

FGNZ began by considering (equation (1) in their paper) the production possibilities set at time t,

$$\mathcal{P}^t = \{(\boldsymbol{x}, \boldsymbol{y}) | \boldsymbol{x} \text{ can produce } \boldsymbol{y} \text{ at time } t\}, \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}_+^p$ and $\boldsymbol{y} \in \mathbb{R}_+^q$ denote vectors of inputs and outputs, respectively. The production possibilities set can be described in terms of its sections

$$\mathcal{Y}^t(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}_+^q \mid (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{P}^t\}, \tag{2}$$

or output correspondence sets. Typical assumptions (*e.g.,* Shephard, 1970; Färe, 1988), which we adopt, include: (i) $\mathcal{P}^t$ is convex, and $\mathcal{Y}^t(\boldsymbol{x})$ is convex, bounded, and closed for all $\boldsymbol{x} \in \mathbb{R}_+^p$; (ii) all production requires the use of some inputs, *i.e.,* $(\boldsymbol{x}, \boldsymbol{y}) \notin \mathcal{P}^t$ if $\boldsymbol{y} \geq 0$, $\boldsymbol{x} = 0$; and (iii) both inputs and outputs are strongly disposable, *i.e.,* if $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{P}^t$ then $\widetilde{\boldsymbol{x}} \geq \boldsymbol{x} \Rightarrow (\widetilde{\boldsymbol{x}}, \boldsymbol{y}) \in \mathcal{P}^t$ and $\widetilde{\boldsymbol{y}} \leq \boldsymbol{y} \Rightarrow (\boldsymbol{x}, \widetilde{\boldsymbol{y}}) \in \mathcal{P}^t$.

The upper boundary of $\mathcal{P}^t$ is sometimes referred to as the *technology* or the *production frontier,* and is given by the intersection of $\mathcal{P}^t$ and the closure of its compliment, or

$$\mathcal{T}^t = \{(\boldsymbol{x}, \boldsymbol{y}) \mid (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{P}^t,\ (\boldsymbol{x}/\lambda, \boldsymbol{y}) \notin \mathcal{P}^t\ \forall\ \lambda > 1,\ (\boldsymbol{x}, \boldsymbol{y}/\theta) \notin \mathcal{P}^t\ \forall\ \theta < 1\}. \tag{3}$$

Let $(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t)$ denote the input and output vectors of production unit i at time t. The Shephard (1970) output distance function for the ith production unit at time t_j, relative to the technology existing at time t_k, is defined as

$$D^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}) \equiv \inf\left\{\theta > 0 \mid (\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}/\theta) \in \mathcal{P}^{t_k}\right\}. \tag{4}$$

The distance function $D^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j})$ gives a normalized measure of distance from the ith production unit's position in the input/output space at time t_j to the boundary of the production set at time t_k in the hyperplane where inputs remain constant. If $t_j = t_k = t$, then we have a measure of efficiency relative to the contemporaneous technology, and $D^t(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t) \leq 1$, with $D^t(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t) = 1$ indicating that the ith production unit is on the boundary of the production set and hence is technically efficient. If $t_j \neq t_k$, then $D^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j})(<,=,>)1$. Note that equation (4) corresponds to equations (2)–(3) in FGNZ.

Now define the set $\mathcal{V}^t$ as the convex cone (with vertex at the origin) spanned by $\mathcal{P}^t$; then $\mathcal{P}^t \subseteq \mathcal{V}^t$. If $\mathcal{T}^t$ exhibits constant returns to scale (CRS) everywhere, then the technology $\mathcal{T}^t$ implies a mapping $\boldsymbol{x} \rightarrow \boldsymbol{y}$ that is homogeneous of degree 1; *i.e.,* $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}^t$ implies $(\lambda\boldsymbol{x}, \lambda\boldsymbol{y}) \in \mathcal{T}^t$ for all $\lambda > 0$. In this case, $\mathcal{P}^t = \mathcal{V}^t$.

If $\mathcal{T}^t$ does not exhibit CRS everywhere, then by construction, $\mathcal{P}^t \subset \mathcal{V}^t$. Analogous to (4), the distance function

$$\Delta^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}) \equiv \inf\left\{\theta > 0 \mid (\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}/\theta) \in \mathcal{V}^{t_k}\right\} \tag{5}$$

gives a normalized measure of distance from the ith production unit's position in the input/output space at time t_j to the boundary of $\mathcal{V}^{t_k}$ in the hyperplane where inputs remain constant. By construction, if $t_j = t_k = t$, then $\Delta^t(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t) \leq D^t(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t) \leq 1$; if $\mathcal{P}^t = \mathcal{V}^t$, then $\Delta^t(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t) = D^t(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t)$ and CRS prevail everywhere on the technology.

FGNZ defined (and decomposed) their Malmquist index (equations 6–7 of their paper) as

$$\begin{aligned}\mathcal{M}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}, \boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) &\equiv \left(\frac{D^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \times \frac{D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{D^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}\right)^{1/2} \\ &= \underbrace{\left(\frac{D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}\right)}_{=\text{EFFCH}} \times \underbrace{\left(\frac{D^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})} \times \frac{D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}{D^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}\right)^{1/2}}_{=\text{TECHCH}}.\end{aligned} \tag{6}$$

Following equation (7) of their paper, FGNZ stated that the term labelled EFFCH "measures the change in relative efficiency (*i.e.,* the change in how far observed production is from maximum potential production)" between times t_1 and t_2, while the term labelled TECHCH "captures the shift in technology between the two periods" evaluated in the hyperplanes containing where the inputs for production unit i are held constant at times t_1

and t_2. Note that FGNZ defined their index in terms of the *true* output distance function defined in (4), which depends on the *true* production possibilities set defined in (1).

Unfortunately, the index defined in the first line of (6) can only measure productivity change if the underlying, true technology exhibits constant returns to scale everywhere (so that the distance functions defined in (4) and (5) are equivalent).[1] Since the shape of the technology can never be known with certainty, this is most disturbing. To see this, consider Figure 1, where production units produce a single output y from a single input x. Suppose the true technology at times t_1 and t_2 is represented by the solid ray emanating from the origin, so that there is no technical change. Suppose a production unit produces at point A at time t_1, and at point B at time t_2. Since the proportion of output to input is the same at A and B, the production unit experiences no change in productivity. At A, the output distance function is $CA/CE < 1$, and at B the output distance function is $DB/DG < 1$. Since $CA/CE = DB/DG$ by construction, there is no change in efficiency. Moreover, the first line of (6) gives

$$\mathcal{M} = \left(\frac{DB/DG}{CA/CE} \times \frac{DB/DG}{CA/CE} \right)^{1/2} = 1, \tag{7}$$

indicating no change in productivity. But now suppose the true technology is given by the curve passing through points E and F, and that again there is no technical change while the production unit moves from A to B. In this example, the first line of (6) gives

$$\mathcal{M} = \left(\frac{DB/DF}{CA/CE} \times \frac{DB/DF}{CA/CE} \right)^{1/2} > 1 \tag{8}$$

(since by construction, $CA/CE < DB/DF$), incorrectly indicating an increase in productivity when there has been none.

The example of Figure 1 illustrates that to measure productivity changes, one must measure distances to a line such as the one passing through points E and G, rather than

[1]FGNZ remarked (p. 71) that they were illustrating (in their Figure 1) the Malmquist index and the decomposition into ΔEff and ΔTech for constant-returns-to-scale technology. But, whether the technology is CRS or otherwise is an empirical question that can only be tested by an appropriate statistical procedure. RD stated in their first paragraph that "the Malmquist productivity index is correctly measured by the ratio of CRS distance functions even when the technology exhibits variable returns to scale," and so seem to recognize the problem.

distances to the technology represented by the curve passing through E and F. In terms of the earlier notation, the Malmquist index can be redefined and decomposed as

$$
\begin{aligned}
\Pi(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}, \boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) &\equiv \left(\frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \times \frac{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)^{1/2} \\
&= \underbrace{\left(\frac{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)}_{=\Delta\text{Eff}} \times \underbrace{\left(\frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})} \times \frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)^{1/2}}_{=\Delta\text{Tech}},
\end{aligned}
\tag{9}
$$

which will correctly measure changes in productivity regardless of the shape of the true technology, since distances are measured relative to the boundary of the conical hull of the production set. This is the Malmquist index that FGNZ estimated, although it is not the one they defined in equations (6)–(7) of their paper.[2] The first line of (9) is also equivalent to the productivity index defined in equation (12) of the RD paper and is one of the quantities that RD attempted to estimate.

The term ΔEff in (9) can be further decomposed, so that the Malmquist index can be written as

$$
\begin{aligned}
\Pi(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}, \boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) = &\underbrace{\left(\frac{D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)}_{=\Delta\text{PureEff}} \times \underbrace{\left(\frac{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})/D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})/D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)}_{=\Delta\text{Scale}} \\
&\times \underbrace{\left(\frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})} \times \frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)^{1/2}}_{=\Delta\text{Tech}},
\end{aligned}
\tag{10}
$$

where $\Delta\text{PureEff} \times \Delta\text{Scale} = \Delta\text{Eff}$. This is analogous to the "enhanced decomposition" of the Malmquist index FGNZ described in footnote 16 of their paper.

To illustrate the decompositions of productivity change in (9) and (10), consider once again the example in Figure 1 where the production unit moves from point A at time t_1

[2] FGN wrote, "The technical change term used in both RD and FGNZ can be summarized as" just before their equation (1); the term in their equation (1) with "j = CRS" would seem to be identical to ΔTech defined here in (9) except for differences in notation. However, FGN defined the technical change measure in their equation (1) in terms of distance functions defined in equations (2)–(3) of FGNZ, which are defined in terms of the production possibilities set defined here in (1) and in equation (1) of FGNZ. Thus, FGN seem to inadvertently imply that the true technology exhibits constant returns to scale, which is not necessarily the case. At other points, FGN acknowledged that the true technology might not exhibit constant returns to scale.

to point B at time t_2, but the technology (given by the curve passing through points E and F) remains unchanged. From (9), $\Pi(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}, \boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) = 1$, $\Delta\text{Eff} = \frac{CA/CE}{DB/DG} = 1$, and $\Delta\text{Tech} = \left(\frac{CA/CE}{CA/CE} \times \frac{DB/DG}{DB/DG}\right)^{1/2} = 1$. This is the desired result for changes in productivity and technology, since there have been no changes in technology or productivity in this example. But, the technical efficiency of the production unit at B is greater than at A; *i.e.,* $\frac{DB}{DF} > \frac{CA}{CE}$. While the first line of (9) gives the correct version of the Malmquist productivity index, the decomposition into ΔEff and ΔTech in the second line is not meaningful. From (10), $\Delta\text{PureEff} = \frac{DB/DF}{CA/CE} > 1$ and $\Delta\text{Scale} = \frac{(DB/DG)/(DB/DF)}{(CA/CE)/(CA/CE)} = \frac{DB/DG}{DB/DF} = \frac{CA/CE}{DB/DF} = 1/\Delta\text{PureEff} < 1$. Pure efficiency increases, since at point B the production unit lies closer to the technology than at point A in the sense of the distance function defined in (4). Scale efficiency decreases, since the projections of A and B onto the technology in the output direction yields points E and F, and at E, the technology exhibits CRS, but at F the technology exhibits decreasing returns to scale. In this example, the increase in pure efficiency is exactly offset by a decrease in scale efficiency. The decomposition of the Malmquist index in (10) appears sensible in the context of Figure 1.

Now consider a slightly different scenario, illustrated in Figure 2. In this scenario, the production unit is located at point B at times t_1 and t_2, but the technology shifts upward everywhere except at point E. In this case, (9) gives $\Delta\text{Eff} = \frac{DB/DG}{DB/DG} = 1$ and $\Delta\text{Tech} = \left(\frac{DB/DG}{DB/DG} \times \frac{DB/DG}{DB/DG}\right)^{1/2} = 1$, so that $\Pi(t_1, t_2) = 1$. Clearly, there has been no change in productivity, but from the perspective of point B, there have been changes in both efficiency and technology. Moreover, from (10), $\Delta\text{PureEff} = \frac{DB/DH}{DB/DF} = \frac{DF}{DH} < 1$ and $\Delta\text{Scale} = \frac{(DB/DG)/(DB/DH)}{(DB/DG)/(DB/DF)} = \frac{DH}{DF} > 1$. Pure efficiency declines, because at time t_2 the production unit is farther from the contemporaneous technology than at time t_1; this decline is exactly offset by an increase in the scale efficiency of the production unit in the sense that when point B is projected onto the two technologies, H is closer to G than is F. But, (10) also yields $\Delta\text{Tech} = 1$, which cannot be correct.[3]

[3]FGN justified the technical change measure based on constant returns to scale in FGNZ by noting (p. 1041), "In FGNZ we think of technical change as change in maximal average product between period t and $t+1$... this is also consistent with the notion of technical change as defined by Robert A. Solow

To remedy this problem, ΔTech can be decomposed to obtain

$$
\begin{aligned}
\Pi(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}, \boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) = {} & \underbrace{\left(\frac{D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}\right)}_{=\Delta\text{PureEff}} \times \underbrace{\left(\frac{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})/D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})/D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}\right)}_{=\Delta\text{Scale}} \\
& \times \underbrace{\left(\frac{D^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})} \times \frac{D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}{D^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}\right)^{1/2}}_{=\Delta\text{PureTech}} \\
& \times \underbrace{\left(\frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})/D^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})/D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})} \times \frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})/D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})/D^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}\right)^{1/2}}_{=\Delta\text{ScaleTech}},
\end{aligned} \tag{11}
$$

where ΔPureTech $\times$ ΔScaleTech replaces ΔTech in (10). The term ΔPureTech is a geometric mean of two ratios. The first ratio measures the shift in the true technology relative to the production unit's position at time t_2, while the second ratio measures the shift in the true technology relative to the production unit's position at time t_1. Following the spirit of the name *pure efficiency change* given to ΔPureEff, we call the term ΔPureTech *pure change in technology.*

The term ΔScaleTech also consists of a geometric mean of two ratios. The denominator in the first of these ratios is a measure of scale efficiency for the production unit at time t_2, and is identical to the numerator in ΔScale. This term measures the distance between the upper boundaries of $\mathcal{P}^{t_2}$ and $\mathcal{V}^{t_2}$ relative to $\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}$ in the hyperplane where $\boldsymbol{x}_i^{t_2}$ is held constant. The term in the numerator of the first ratio in ΔScaleTech is similar, but measures the distance between $\mathcal{P}^{t_1}$ and $\mathcal{V}^{t_1}$ relative to the same point, namely $\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}$. Thus, the ratio measures the change in the scale, or shape, *of the technology* between times t_1 and t_2, relative to the ith production unit's location at time t_2. Similar reasoning reveals

(1957)." But Solow used time-series data from a single country, and had little choice but to assume constant returns to scale. FGNZ and RD, on the other hand, examined panel data on 17 countries with economies of widely varying size. While one can certainly define technical change as FGN have, it is not clear that changes in maximal average product are relevant for small economies which may be operating under the increasing-returns-to-scale part of a variable-returns-to-scale technology, far from the part of the technology that determines maximal average product. The computational problems that RD encountered and which FGN refer to at several points involve infeasible constraints on the linear programs used to compute distance function estimates; the problem is perhaps better viewed as an identification problem. For example, RD's Table 1, the technical change index for Ireland is not identified by the data.

that the second ratio in ΔScaleTech measures the change in the shape of the technology between times t_1 and t_2 relative to the ith production unit's location at time t_1. Thus ΔScaleTech describes the *change in the scale of the technology* at two fixed points (defined by the production unit's location at times t_1 and t_2). Any changes indicated by ΔScale may be caused either by (i) changes in the shape of the technology, (ii) changes in the location of the production unit in input/output space between t_1 and t_2, or (iii) some combination of (i) and (ii); ΔScale reflects changes in the scale efficiency *of the production unit.* However, any changes indicated by ΔScaleTech can only be due to changes in the shape of the technology, since the reference points are fixed.

To illustrate these points, return to the example illustrated in Figure 2. From (11),

$$\Delta\text{PureTech} = \left(\frac{DB/DF}{DB/DH} \times \frac{DB/DF}{DB/DH}\right)^{1/2} = \frac{DH}{DF} > 1$$

and

$$\Delta\text{ScaleTech} = \left(\frac{(DB/DG)/(DB/DF)}{(DB/DG)/(DB/DH)} \times \frac{(DB/DG)/(DB/DF)}{(DB/DG)/(DB/DH)}\right)^{1/2} = \frac{DF}{DH} < 1.$$

So, while there is an increase in pure technology from the persepective of the production unit at point B, this is exactly offset by a change in the scale of the technology, and productivity remains unchanged, as must be the case if the production unit does not move. This example also reveals that values ΔScaleTech < 1 indicate a flattening of the technology, while ΔScaleTech > 1 indicates increasing curvature, or a change away from CRS.

RD proposed estimators of ΔPureEff, ΔPureTech, and a term they denoted SCH which equals the product ΔScale $\times$ ΔScaleTech:[4]

$$\text{SCH} = \left(\frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})/D^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})/D^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \times \underbrace{\frac{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})/D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})/D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}}_{=\Delta\text{Scale}}\right)^{1/2}. \qquad (12)$$

[4]RD did not call the various expressions in their paper *estimators,* but they illustrated them in terms of an empirical example (their Figure 1) and used them to compute the estimates that appear in their Table 1.

This term is a geometric mean of two ratios; the second ratio in parentheses on the right-hand side of (12) is the ΔScale term defined previously, which measures changes in the returns to scale faced by the production unit. In the first ratio inside the parentheses, the numerator measures scale efficiency the production would have experienced at time t_1 had it been located at its position at time t_2, while the denominator measures the scale efficiency the production unit would have experienced at time t_2 had it remained at its location at time t_2. For the scenario in Figure 2, we have SCH = 1 (since ΔScale = 1/ΔScaleTech), indicating no change. But, as we have seen, two things have changed: The scale efficiency of the production unit at B, and the scale of the technology; in this simple illustration, the two changes have exactly offsetting effects, but this would not likely be the case in real applications. In a more general setting where we might obtain SCH $\neq 1$, it would be impossible to know from RD's approach whether the change was due to a movement of the production unit, a change in the shape of the technology, or some combination of these two things. RD offered scant explanation of their term SCH; they merely stated at the end of the paragraph containing their equation (16) that

> "Our measure [SCH] is a geometric mean of the ratios of scale efficiencies of the two [input/output] bundles using in turn the VRS [*i.e.*, variable returns to scale] technologies from the two periods as the benchmark. In that sense, it is more in the spirit of a Fisher index."

While SCH may be in the spirit of a Fisher index, this is of little consequence since the measure necessarily confuses changes in the scale efficiency experienced by the production unit and changes in the scale of the technology.[5]

[5] RD rearranged terms in (12) to write SCH as

$$\text{SCH} = \left(\frac{\Delta^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})/D^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\Delta^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})/D^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})} \times \frac{\Delta^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})/D^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}{\Delta^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})/D^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)^{1/2}.$$

The first ratio inside the parentheses would measure the change in scale efficiency experienced by the production unit between times t_1 and t_2 if its location remained fixed at $(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})$. Similarly, the second ratio inside the parentheses would measure the change in scale efficiency experienced by the production unit between times t_1 and t_2 if its location remained fixed at $(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})$. But, production units' locations typically do not remain fixed over time. FGN recognized that SCH "may incorrectly identify the scale properties of the underlying technology," and illustrated the point with an example in their footnote 7. But, FGN did not distinguish between changes in returns to scale faced by individual production units and

II. Estimation Issues

Unfortunately, the production set $\mathcal{P}^t$, the conical hull $\mathcal{V}^t$, and the output correspondence sets $\mathcal{Y}^t(\boldsymbol{x})$ are unobserved; therefore, the distance functions in (4)–(5) are also unobserved and must be estimated. Since the Malmquist index defined in the first line of (9) and its decomposition in (11) depends on the true distance function values defined in (4)–(5), the Malmquist index as well as its components must also be estimated. Neither FGNZ nor RD discussed this; implicitly, both assumed the true quantities could be determined from data, and that there is no difference between the underlying, true quantities and their estimates. For example, RD stated (in the last paragraph of their section I) that

> "The own and cross-period output-oriented distance functions can be obtained by solving the appropriate linear programming problems specified by [FGNZ] (their problem (17) on page 75)."

FGNZ made similar remarks; *e.g.,* leading to equations (16) and (17) in their paper, FGNZ stated that they were "computing" the distance functions defined here in (4)–(5) by solving linear programming problems. In fact, both papers solved linear programs to *estimate* the distance functions defined in (4)–(5).[6]

Before meaningful inferences can be drawn from any empirical results that might be obtained, one must define an econometric model in order to obtain some understanding of the properties of the estimators being used. We adopt an econometric model similar to that proposed by Kneip *et al.* (1996), which can be summarized by the following assumptions on the data-generating process:

Assumption A1: *For a given time t, $(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t)$, $i = 1, \ldots, n$ are identically, independently distributed (iid) random variables on $\mathcal{P}^t$.*

changes in the shape of the technology, and they ignored the fact that SCH confuses these two different phenomena.

[6] After introducing an estimator of the distance function $\Delta^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j})$ without defining the quantity they wish to estimate (and proceeding as if estimates yielded by the estimator were *true* values), RD contributed further confusion when they refered to this estimator and the corresponding estimates as a "CRS distance function," and in addition made reference to "the CRS production possibility set" in the text following equation (3) in their paper. Later they recognized that some points in this set may not be feasible if the technology does not exhibit CRS everywhere. But, in this case, what they refered to as "the CRS production possibility set" cannot be a production possibility set if it contains infeasible points.

Assumption A2: *The density $f^t(\boldsymbol{x})$ of $\boldsymbol{x}$ at time t has compact support $\mathcal{D}^t \subseteq \mathbb{R}^p_+$, and for all $\boldsymbol{x} \in \mathcal{D}^t$, $\boldsymbol{y}$ has conditional density $f^t(\boldsymbol{y} \mid \boldsymbol{x})$ on $\mathcal{Y}^t(\boldsymbol{x})$.*

Assumption A3: *There exist constants $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ such that for all $\boldsymbol{x} \in \mathcal{D}^t$ and all $\boldsymbol{y} \in \mathcal{Y}^t(\boldsymbol{x})$ and $\omega \in [1, 1+\varepsilon_2]$ such that $(\boldsymbol{x}, \omega\boldsymbol{y}) \in \mathcal{T}^t$, $f^t(\boldsymbol{y}|\boldsymbol{x}) \geq \varepsilon_1$.*

Assumption A4: *$D^t(\boldsymbol{x}, \boldsymbol{y})$ is differentiable in both its arguments.*

Assumption A3 is required to prove statistical consistency for estimators of $D^t(\boldsymbol{x}, \boldsymbol{y})$, since it ensures that when the sample size increases we will observe production units $(\boldsymbol{x}, \boldsymbol{y})$ near the true frontier. Assumption A4 is actually stronger than what is required to prove consistency, but avoids complications in deriving the rates of convergence (see Kneip *et al.* for details). The joint density $f^t(\boldsymbol{x}, \boldsymbol{y})$ represents the data generating process at time t; estimation of the distance function $D^t(\boldsymbol{x}, \boldsymbol{y})$ amounts to estimating the boundary of the support of the conditional density $f^t(\boldsymbol{y} \mid \boldsymbol{x})$.

Given an estimator $\widehat{\mathcal{P}}^t$ of $\mathcal{P}^t$, the distance function $D^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j})$ defined in (4) may be estimated by

$$\widehat{D}^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}) \equiv \inf\left\{\theta > 0 \mid (\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}/\theta) \in \widehat{\mathcal{P}}^{t_k}\right\}. \tag{13}$$

Similarly, given an estimator $\widehat{\mathcal{V}}^t$ of $\mathcal{V}^t$, the distance function $\Delta^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j})$ defined in (5) may be estimated by

$$\widehat{\Delta}^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}) \equiv \inf\left\{\theta > 0 \mid (\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}/\theta) \in \widehat{\mathcal{V}}^{t_k}\right\}. \tag{14}$$

Several estimators of $\mathcal{P}^t$ are possible. Given a sample of n observations of inputs and outputs of production units at time t denoted by

$$\mathcal{S}_n^t = \{(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t) \mid i = 1, \ldots, n\}, \tag{15}$$

the union of the convex and free disposal hulls of the points in $\mathcal{S}_n^t$ is given by

$$\widehat{\mathcal{P}}_n^t = \{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{y} \leq \boldsymbol{Y}^t\boldsymbol{\tau},\ \boldsymbol{x} \geq \boldsymbol{X}^t\boldsymbol{\tau},\ \boldsymbol{i}\boldsymbol{\tau} = 1,\ \boldsymbol{\tau} \in \mathbb{R}^n_+\}, \tag{16}$$

where $\boldsymbol{Y}^t = [\boldsymbol{y}_1^t\ \ldots\ \boldsymbol{y}_n^t]$, $\boldsymbol{X}^t = [\boldsymbol{x}_1^t\ \ldots\ \boldsymbol{x}_n^t]$, with each $\boldsymbol{x}_i^t$, $\boldsymbol{y}_i^t$ $i = 1, \ldots, n$ denoting the $(p \times 1)$ and $(q \times 1)$ vectors of observed inputs and outputs, respectively, and where $\boldsymbol{i}$ is a $(1 \times n)$ vector of ones and $\boldsymbol{\tau} = [\tau_1 \ \ldots \ \tau_n]'$ is a $(n \times 1)$ vector of intensity variables. An estimator of the distance function defined by (4) may be obtained by substituting the estimator $\widehat{\mathcal{P}}_n$ for $\mathcal{P}$ to obtain

$$\widehat{D}_n^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}) = \max\left\{\theta \mid \theta \boldsymbol{y}_i^{t_j} \leq \boldsymbol{Y}_i^{t_k}\boldsymbol{\tau},\ \boldsymbol{x}_i^{t_j} \geq \boldsymbol{X}_i^{t_k}\boldsymbol{\tau},\ \boldsymbol{i\tau} = 1,\ \boldsymbol{\tau} \in \mathbb{R}_+^n\right\}. \tag{17}$$

Korostelev *et al.* (1995) proved that $\widehat{\mathcal{P}}_n^t$ is a consistent estimator of $\mathcal{P}^t$ and give rates of convergence. Kneip *et al.* (1996) proved that under assumptions A1–A4, $\widehat{D}_n(\boldsymbol{x}, \boldsymbol{y}) - D(\boldsymbol{x}, \boldsymbol{y}) = O_p(n^{-\frac{2}{p+q+1}})$. However, this is only an asymptotic result. In the application considered by FGNZ and by RD, although we asymptotically have root-n convergence since we have $p + q = 3$, we also have $n = 17$, and performance of the estimator in small samples remains unknown.

The convex polyhedral cone (with vertex at the origin) spanned by $\mathcal{S}_n^t$ (or equivalently, by $\widehat{\mathcal{P}}_n^t$),

$$\widehat{\mathcal{V}}_n^t = \{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{y} \leq \boldsymbol{Y}^t\boldsymbol{\tau},\ \boldsymbol{x} \geq \boldsymbol{X}^t\boldsymbol{\tau},\ \boldsymbol{\tau} \in \mathbb{R}_+^n\}, \tag{18}$$

gives an estimator of $\mathcal{V}$. Substituting $\widehat{\mathcal{V}}_n^t$ for $\mathcal{V}^t$ in (5) to obtain

$$\widehat{\Delta}_n^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j}) = \max\left\{\theta \mid \theta \boldsymbol{y}_i^{t_j} \leq \boldsymbol{Y}^{t_k}\boldsymbol{\tau},\ \boldsymbol{x}_i^{t_j} \geq \boldsymbol{X}^{t_k}\boldsymbol{\tau},\ \boldsymbol{\tau} \in \mathbb{R}_+^n\right\} \tag{19}$$

yields an estimator of $\Delta^{t_k}(\boldsymbol{x}_i^{t_j}, \boldsymbol{y}_i^{t_j})$. It is straightforward to extend the results of Korostelev *et al.* (1995) to prove that $\widehat{\mathcal{V}}_n^t$ is a consistent estimator of $\mathcal{V}^t$ (and therefore a consistent estimator of $\mathcal{P}^t$ if $\mathcal{P}^t = \mathcal{V}^t$, *i.e.,* if $\mathcal{T}^t$ exhibits CRS everywhere) and that $\widehat{\Delta}_n(\boldsymbol{x}, \boldsymbol{y})$ is a consistent estimator of $\Delta_n(\boldsymbol{x}, \boldsymbol{y})$. By construction,

$$\widehat{\mathcal{P}}_n^t \subseteq \widehat{\mathcal{V}}_n^t. \tag{20}$$

If $\mathcal{P}_n^t$ is convex, and $\mathcal{T}^t$ does not exhibit CRS everywhere, then $\widehat{\mathcal{P}}^t$ is a consistent estimator of $\mathcal{P}^t$, but $\widehat{\mathcal{V}}_n^t$ is not. We say that $\widehat{\mathcal{V}}_n^t$ is constrained relative to $\widehat{\mathcal{P}}_n^t$; *i.e.,* $\widehat{\mathcal{V}}_n^t$ is a more restrictive estimator of $\mathcal{P}^t$ than $\widehat{\mathcal{P}}_n^t$ in the sense that $\widehat{\mathcal{P}}_n^t$ will be consistent regardless of the shape of the convex set $\mathcal{P}^t$, while $\widehat{\mathcal{V}}_n^t$ will be consistent only if $\mathcal{T}^t$ exhibits constant returns to scale everywhere, *i.e.,* if $\mathcal{P}^t = \mathcal{V}^t$.

FGN recognized the fact represented in (20) in the third paragraph of their paper. But, they went on to say that the true technology is bounded by the upper boundaries of $\widehat{\mathcal{P}}_n^t$ and $\widehat{\mathcal{V}}_n^t$. This cannot be true for any finite sample since $\widehat{\mathcal{P}}_n^t$ and $\widehat{\mathcal{V}}_n^t$ are biased estimators, as discussed above. FGN's statement can only be true when $n = \infty$, but in this case the statement is trivial, since $\widehat{\mathcal{P}}_\infty^t$ would correspond with the true production set $\mathcal{P}^t$ given that $\widehat{\mathcal{P}}_n^t$ consistently estimates $\mathcal{P}^t$.

Replacing the true distance function values in the first line of (9) and in (11) with their corresponding estimators yields estimators of the Malmquist index and its components:

$$
\begin{aligned}
\widehat{\Pi}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}, \boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) \equiv & \left(\frac{\widehat{\Delta}^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\widehat{\Delta}^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \times \frac{\widehat{\Delta}^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\widehat{\Delta}^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)^{1/2} \\
= & \underbrace{\left(\frac{\widehat{D}^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\widehat{D}^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)}_{=\widehat{\Delta \mathrm{PureEff}}} \times \underbrace{\left(\frac{\widehat{\Delta}^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) / \widehat{D}^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\widehat{\Delta}^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}) / \widehat{D}^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)}_{=\widehat{\Delta \mathrm{Scale}}} \\
& \times \underbrace{\left(\frac{\widehat{D}^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\widehat{D}^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})} \times \frac{\widehat{D}^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}{\widehat{D}^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)^{1/2}}_{=\widehat{\Delta \mathrm{PureTech}}} \\
& \times \underbrace{\left(\frac{\widehat{\Delta}^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) / \widehat{D}^{t_1}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})}{\widehat{\Delta}^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2}) / \widehat{D}^{t_2}(\boldsymbol{x}_i^{t_2}, \boldsymbol{y}_i^{t_2})} \times \frac{\widehat{\Delta}^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}) / \widehat{D}^{t_1}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})}{\widehat{\Delta}^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1}) / \widehat{D}^{t_2}(\boldsymbol{x}_i^{t_1}, \boldsymbol{y}_i^{t_1})} \right)^{1/2}}_{=\widehat{\Delta \mathrm{ScaleTech}}} .
\end{aligned}
\tag{21}
$$

It is apparently quite difficult, if not impossible, to derive analytic results on the sampling properties of individual distance function estimators.[7] This is even more true for functions of several distance function estimators as in (21). Fortunately, Efron's (1979) bootstrap ideas can be adapted to problems such as the one here.

III. Empirical Analysis

Simar and Wilson (1997a) developed a smooth bootstrap procedure which may be used to estimate confidence intervals for distance functions such as those defined in (4)–(5).

[7]Gijbels *et al.* (1996) derived asymptotic results for $\widehat{D}^t(\boldsymbol{x}, \boldsymbol{y})$ for the special case of one input and one output.

These methods were extended in Simar and Wilson (1997b) to estimate confidence intervals for the Malmquist index and its various components. The intuitive idea of the bootstrap procedure is to replicate the data-generating process in a way that yields an empirical approximation of the sampling distribution of the statistic of interest. This can then be used to estimate confidence intervals for the parameter being estimated. The typical, naive bootstrap involves resampling from an observed empirical distribution of the data, but in the present case such an approach leads to inconsistent bootstrap estimates (see Simar and Wilson, 1997c, for discussion and a simple demonstration of this point). The problem results from the fact that the empirical distribution of the data does not yield a useful approximation of the underlying data-generating process near the boundary of support of $f(\boldsymbol{y} \mid \boldsymbol{x})$. The problem is alleviated by smoothing, wherein the bootstrap resampling is accomplished by drawing from a consistent, nonparametric estimate of $f(\boldsymbol{y} \mid \boldsymbol{x})$. In Simar and Wilson (1997b), draws are taken from a nonparametric, kernel estimate of the bivariate density $f(\boldsymbol{y}^{t_1}, \boldsymbol{y}^{t_2} \mid \boldsymbol{x}^{t_1}, \boldsymbol{x}^{t_2})$ to account for any temporal correlation.

It is straightforward to adapt the bootstrap procedure in Simar and Wilson (1997b) to the present case. We use the same data that RD used, *i.e.,* annual data for 1979–1990 on labor, capital, and real gross domestic product for 17 OECD countries. The data were taken from the Penn World Tables (version 5.6) described by Summers and Heston (1991). These data constitute an update of the data used by FGNZ, and include data on two additional years that were unavailable when FGNZ wrote their paper.

All computations were performed using Fortran code written by the authors. Solutions to linear programs were verified by comparing our results with results obtained using independently coded routines in Matlab, as well as results from the SAS LP procedure (SAS, 1989). Table 1 shows geometric means of the various indices over the 12 periods 1979–80, 1980–81, ..., 1989–90, by country. The last row of the table shows geometric means over the 12 periods and the 17 countries. Asterisks are used to indicate means that are significantly different from unity, determined by whether the estimated bootstrap confidence intervals contained unity.[8]

The first column of numbers in Table 1 contains estimates of average annual productivity change. Twelve of the estimates (71 percent) are significantly different from unity. Of these, 10 indicate increasing productivity; only Austria and Italy show a mean decline in productivity, and these are small (0.53 and 0.05 percent per annum, respectively).[9] Our estimates of mean productivity change are close to the values reported by FGNZ, but very different from those reported by RD. Whereas FGNZ reported annual average productivity growth of 0.85 percent for the US, we estimate the growth rate at 0.21 percent per year. As noted in the introduction, RD reported a 5.56 percent average annual *decline* in productivity for the US, purportedly using the same data and the same estimation method. In addition, our estimates of average annual productivity growth for Australia, France, Ireland, and Sweden, as well as our estimates for the US and for the entire sample, are significantly different from unity and *in the opposite direction* compared to the values reported by RD.[10] Our estimate of 1.0067 mean annual productivity change for all countries is insignificantly different from the value of 1.0070 reported by FGNZ. However, our estimate of 1.0021 for mean annual productivity change for the US over 1979–90 is significantly different at the .1 level from the value of 1.0085 reported by FGNZ for the US over 1979–88.

The next four columns in Table 1 show estimates average annual changes in pure efficiency, scale efficiency, in pure technology, and scale of technology, corresponding to the decompostion of productivity change represented by (21). Only four of these values are significantly different from unity. The last two columns of Table 1 show mean estimates for

[8]The Simar and Wilson (1997b) bootstrap methodology was used to estimate confidence intervals for country and period-specific values, which we do not report here to conserve space (these are available from the authors on request). When averaging country-specific estimates over time, we also average the corresponding bootstrap values over time to obtain estimates of significance used to determine where the asterisks belong in Table 1. This contrasts with the methodology of Atkinson and Wilson (1995) which conditions on the original estimates of productivity change, *etc.*

[9]Subtracting 1 from the number reported in the table gives the average increase or decrease per annum for the relevant performance estimate.

[10]Our estimates of average annual annual productivity growth for Greece and the United Kingdom are also in the opposite direction compared to the corresponding values reported by RD, but insignificantly different from unity.

ΔTech appearing in FGNZ's decomposition and in (10) and for SCH defined by RD and in (12). Recalling that ΔTech was the index used by FGNZ to measure change in technology, the mean estimates of this term shown in Table 1 should agree with the values in the third column of RD's Table 2, but they do not. Moreover, the mean estimates for SCH that we obtained do not correspond to the values reported in the last column of RD's Table 1, as they should; nor do our results for pure technology change agree with the values RD reported in the third column of their Table 1.[11]

None of the mean estimates for ΔTech we report in Table 1 are significantly different from unity. However, this is of little consequence, for as we demonstrated earlier, this term does not provide a useful measure of technology change in general. Mean estimates for RD's SCH term are significant in 12 of 17 cases, but this too is of little consequence since SCH confuses changes in scale efficiency experienced by the production unit and changes in the shape of the technology. We include these estimates only for comparison purposes.

IV. Conclusions

Rolf Färe and Shawna Grosskopf, in a series of papers with various coauthors, have provided useful tools for estimating various notions of productive efficiency as well as changes in productivity. It has only recently been recognized that the linear programming methods they have used in these papers can be given a statistical interpretation. Without a statistical interpretation, it is not meaningful to draw inferences from results obtained with these methods as it is otherwise impossible to know whether the numbers reflect real economic phenomena or merely sampling variation.

What do the Penn World Table data tell us about productivity change in the 17 OECD

[11] Amont the results reported by RD, only our results for ΔPureEff agree with values reported by RD. Our results for ΔScale agree with the results reported by RD except in the case of Japan, where they reported an average of 1.0082, as opposed to our value of 1.0028; presumably, this is a typographical error. However, these indices are the only ones which do not contain distance function estimators which measure distance from an observed input/output combination in one period to an estimate of the technology in a different period. Thus, RD's computer code (they did not report how they computed the values they reported in their tables) apparently gave incorrect solutions for the linear programs used to compute cross-period distance function estimates. RD reported infeasible solutions for their technical change and scale efficiency change indices in the case of Ireland; we found infeasible solutions for some of the cross-period distance functions for Ireland, too.

countries examined by FGNZ and by RD? Table 1 contains the answer to this question; most countries experienced, on average, increasing productivity over 1979-90. Much less can be said about the sources of these changes, however, a fact that is not obvious from reading FGNZ or RD.

REFERENCES

Atkinson, Scott E. and Wilson, Paul W. "Comparing Mean Efficiency and Productivity Scores from Small Samples: A Bootstrap Methodology," *Journal of Productivity Analysis,* July 1995, 6(2), pp. 137–152.

Chavas, Jean-Paul and Cox, Thomas L. "A Non-Parametric Analysis of Productivity: The Case of US and Japanese Manufacturing," *American Economic Review,* June 1990, 80(3), pp. 450–64.

Efron, Bradley. "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics,* March 1979, 7(1), pp. 1–16.

Färe, Rolf. *Fundamentals of Production Theory.* Berlin: Springer Verlag, 1988.

Färe, Rolf; Grosskopf, Shawna; Norris, Mary and Zhang, Zhongyang. "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries." *American Economic Review,* March 1994, 84(1), pp. 66–83.

Färe, Rolf; Grosskopf, Shawna; Norris, Mary. "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries: Reply." *American Economic Review,* December 1997, 87(6), pp. 1040–1043.

Gijbels, Irene; Mammen, Enno; Park, Byeong U. and Simar, Léopold. "On Estimation of Monotone and Concave Frontier Functions," Discussion Paper #9611, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 1996.

Kneip, Alois; Park, Byeong U. and Simar, Léopold. "A Note on the Convergence of Nonparametric DEA Efficiency Measures," Discussion Paper #9603, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 1996.

Korostelev, Alexander P.; Simar, Léopold and Tsybakov, Alexander B. (1995), "On Estimation of Monotone and Convex Boundaries, *Publications de l'Institut de Statistique de l'Université de Paris* 1995, 39, pp. 3–18.

Ray, Subhash and Desli, Evangelia. "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries: Comment" *American Economic Review,* December 1997, 87(6), pp. 1033-1039.

SAS Institute, Inc. *SAS/OR User's Guide, Version 6, First Edition.* Cary, NC: SAS Institute, Inc., 1989.

Simar, Léopold and Wilson, Paul W. "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models, *Management Science,* 1997a, in press.

Simar, Léopold and Wilson, Paul W. "Estimating and Bootstrapping Malmquist Indices, *European Journal of Operational Research,* 1997b, forthcoming.

Simar, Léopold and Wilson, Paul W. "Some Problems with the Ferrier/Hirschberg Bootstrap Idea, *Journal of Productivity Analysis,* 1997c, forthcoming.

Shephard, Ronald W. *Theory of Cost and Production Functions.* Princeton: Princeton University Press, 1970.

Solow, Robert M. "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics,* August 1957, 39(3), pp. 312–320.

Summers, Robert and Heston, Alan. "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1987." *Quarterly Journal of Economics,* May 1991, 106, pp. 1–41.

Table 1—Malmquist Productivity Index and its Decomposition; Geometric Means (1979–1990)

Country	$\widehat{\Pi}$	$\widehat{\Delta\text{PureEff}}$	$\widehat{\Delta\text{Scale}}$	$\widehat{\Delta\text{PureTech}}$	$\widehat{\Delta\text{ScaleTech}}$	$\widehat{\Delta\text{Tech}}$	$\widehat{\text{SCH}}$
Australia	1.0112**	0.9990	1.0017	1.0094	1.0010	1.0104	1.0028**
Austria	0.9947**	0.9995	0.9962	1.0085**	0.9905**	0.9989	0.9868**
Belgium	1.0153**	1.0030	1.0019	1.0096*	1.0008	1.0105	1.0027**
Canada	1.0150**	1.0036	1.0009	1.0104	1.0001	1.0105	1.0010**
Denmark	1.0033	0.9976	1.0068	1.0101*	0.9890	0.9990	0.9957**
Finland	1.0257**	1.0107	1.0043	1.0075	1.0029	1.0105	1.0073**
France	1.0115**	1.0011	1.0002	1.0100	1.0001	1.0101	1.0003**
Germany	1.0072**	0.9966	1.0002	1.0105	0.9999	1.0104	1.0002**
Greece	0.9982	0.9981	1.0015	0.9991	0.9996	0.9987	1.0010
Ireland	1.0066**	1.0000	1.0066	—	—	1.0000	—
Italy	0.9995*	1.0048	1.0001	0.9949	0.9997	0.9946	0.9998
Japan	1.0010	1.0003	1.0028	0.9965	1.0014	0.9979	1.0042**
Norway	1.0154**	1.0000	1.0049	1.0108	0.9997	1.0105	1.0046**
Spain	0.9967	0.9970	1.0003	0.9996	0.9997	0.9993	1.0000
Sweden	1.0129**	1.0000	1.0025	1.0073	1.0031	1.0104	1.0056**
UK	0.9982	1.0000	1.0000	0.9982	1.0000	0.9982	1.0000
USA	1.0021*	1.0000	1.0000	1.0058	0.9963	1.0021	0.9963
All	1.0067**	1.0007	1.0018	1.0055	0.9990	1.0042	1.0005

Note: Double asterisk (**) denotes significant difference from unity at .05; single asterisk (*) denotes significant difference from unity at .1.

Figure 1

Firm Moves from A to B, but Technology Does not Change

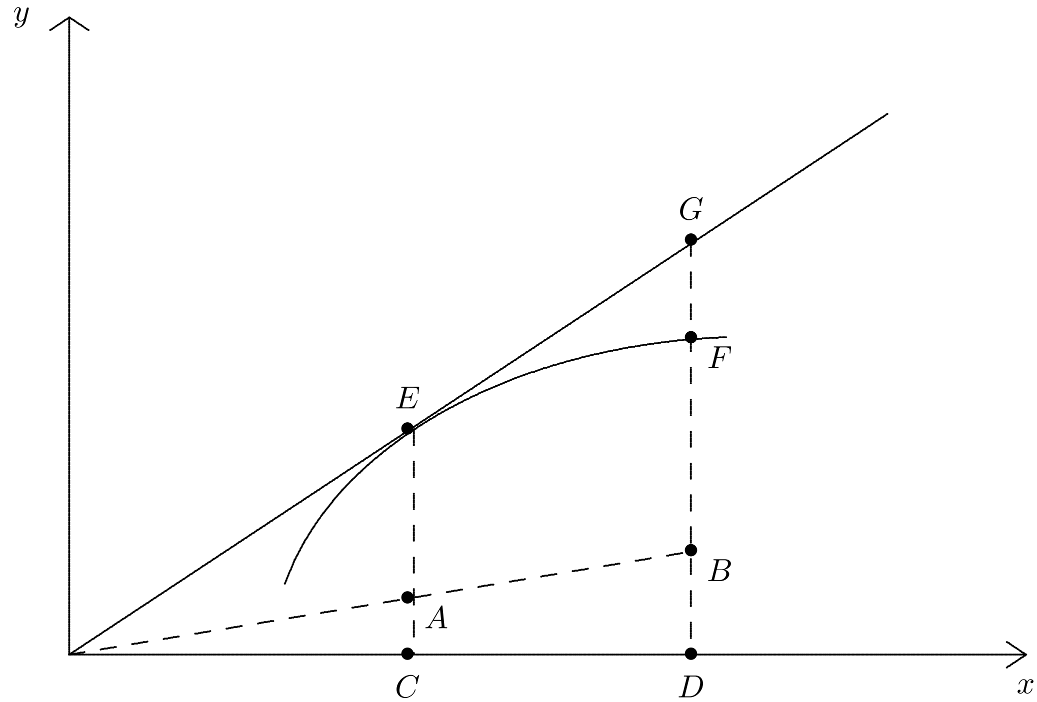

Figure 2

Production Unit Remains at B, but Technology Changes