Economics Bulletin

Appendix and Supplemental material not intended for publication-Round 3

Submission Number:EB-16-00160

Web Appendix

Research in Quantitative Economics of Development Laboratory (LAREQUAD). E-mail : khthabet@yahoo.com. I want to thank John Conley, Baccouche Rafik and the two anonymous referees of Economics Bulletin for their comments and suggestions.
Submitted: Feb 21 2016. **Revised:** February 07, 2018.

Web Appendix to Determinants of total factor productivity growth of Tunisian manufacturing firms

Linear programming techniques to estimate the distances functions

For both periods t^1 and t^2 the production set, and consequently all distances defined from it are unknown. Following Färe et al. (1992), the four distances which make up a Malmquist index can be estimated via linear programming techniques. The distance functions cannot be computed without knowing the frontier production set. In the general case that includes variable returns to scale (VRS), the output-oriented measure of technical efficiency is derived from the following optimization VRS DEA (Data Envelopment Analysis) problem for each firm i (see, for example, Fa¨re et al. 1994, p. 75; Coelli et al. 1998, p. 158):

$$\begin{aligned} & \left[\hat{d}_i^{t_1}\left(x_i^{t_1}, y_i^{t_1}\right)\right]^{-1} \max_{\phi,\lambda} \phi \\ s.c \quad & -\phi y_{ip}^{t_1} + \sum_{j=1}^{n} \lambda_j^{t_1} y_{jp}^{t_1} \geq 0, \qquad p = 1, \ldots, P, \\ & x_{ik}^{t_1} - \sum_{j=1}^{n} \lambda_j^{t_1} x_{jk}^{t_1} \geq 0, \qquad k = 1, \ldots, K, \\ & \sum_{i=1}^{n} \lambda_i^{t_1} = 1 \\ & \lambda_i^{t_1} \geq 0 \qquad i = 1, \ldots, n. \end{aligned} \tag{1}$$

where $x_{ik}^{t_1}$ is the k-th input and $y_{ip}^{t_1}$ is the p-th output for DMUi in period t_1, and $\lambda^{t_1} = (\lambda_1^{t_1}, \ldots, \lambda_n^{t_1})'$ is a vector of weights that forms a convex combination of observed firms relative to which the subject firm's efficiency is evaluated. The convexity constraint ($\sum_{i=1}^{n} \lambda_i = 1$) ensures that an inefficient production unit is ''benchmarked'' against production units of a similar size. The output-oriented measure of technical efficiency of the i-th production unit is given by $\hat{d}_i^{t_1}\left(x_i^{t_1}, y_i^{t_1}\right) = TE_i = {}^{1}/_{\phi_i^*}$. The efficiency determines the amount by which produced output can be proportionally increased, while still using the given inputs level. TE_i varies between zero and one ($0 < TE_i \leq 1$, where $TE_i = 1$ means that the ith production unit is fully efficient and operates on the best-practice frontier).Using t_2 instead of t_1 for the above model, we get $\hat{d}_i^{t_2}\left(x_i^{t_2}, y_i^{t_2}\right)$, the technical efficiency score for DMUi in time period t_2.

Two further linear programming models are needed to estimate the mixed-period cases. The first of the mixed period measures, which is defined as $\hat{d}_i^{t_1}(x_i^{t_2}, y_i^{t_2})$ for each DMUi, is computed as the optimal value to the following linear programming problem:

$$
\begin{aligned}
& \left[\hat{d}_i^{t_1}(x_i^{t_2}, y_i^{t_2})\right]^{-1} \max_{\phi,\lambda} \phi \\
s.c \quad & -\phi y_{ip}^{t_2} + \sum_{j=1}^{n} \lambda_j^{t_1} y_{jp}^{t_1} \geq 0, \qquad p = 1, \ldots, P, \\
& x_{ik}^{2} - \sum_{j=1}^{n} \lambda_j^{t_1} x_{jk}^{t_1} \geq 0, \qquad k = 1, \ldots, K, \\
& \sum_{i=1}^{n} \lambda_j^{t_1} = 1, \\
& \lambda_i^{t_1} \geq 0, \qquad i = 1, \ldots, n.
\end{aligned}
\tag{2}
$$

As Färe et al. (1994) state, observations involved in (2) are from both period t_2 and t_1. The reference technology relative to which $(x_i^{t_2}, y_i^{t_2})$ is evaluated, is constructed from observations in t_1. To compute the second mixed-period distance function, $\hat{d}_i^{t_2}(x_i^{t_1}, y_i^{t_1})$, which is needed in the computation of the output-oriented MPI, t_1 and t_2 superscripts in (2) must simply be reversed.

Bootstrapping Malmquist productivity indices

Bootstrapping involves replicating the data generating process, generating an appropriately large number B of pseudo samples $\tilde{\varphi} = \{(\tilde{x}_{it}, \tilde{y}_{it}); i = 1, \ldots n; t = 1,2\}$ and applying the original estimators to these pseudo samples. For each bootstrap replication b=1; . . . ; B, we use Eq. (1) to measure the distance from each observation in the original sample φ to the frontiers estimated for either period from the pseudodata in $\tilde{\varphi}$. We use the smoothed bootstrap procedure of Simar and Wilson (1999) to draw bootstrap samples of {y,x} and use them to conduct bootstrap inference[1]. In the case of panel data (our case), to control any temporal correlation present in the data[2], Simar and Wilson (1999) use kernel methods to estimate the joint density of $\left\{\left(\hat{d}_i^{t_1}(x_i^{t_1}, y_i^{t_1}) \;\; \hat{d}_i^{t_2}(x_i^{t_2}, y_i^{t_2})\right)\right\}_{i=1}^{n}$

The first step consist to compute the Malmquist productivity index $\widehat{M}(x_i^{t_2}, y_i^{t_2}, x_i^{t_1}, y_i^{t_1})$ for each DMU by solving the linear programming models (1 and 2) and their reversals for each

[1] When bootstrapping distance function estimates from a single cross-section of data, we can use simply a univariate kernel estimator of the density of the original distance function estimates, and then drawing from this estimated density to construct the pseudo samples $\tilde{\varphi}$ as in and Simar and Wilson (1998).

[2] For example, an inefficient firm in period one may be more likely to be inefficient in period two than a firm that is relatively more efficient in period one.

firm (i=1,..., N) in each time (t_1 and t_2) as described in Faire et al (1992, 1995). For simplicity we note $M(t_1, t_2)$ instead $M\left(x_i^{t_2}, y_i^{t_2}, x_i^{t_1}, y_i^{t_1}\right)$.

The procedure consist First, to form $(N \times 1)$ vectors A and B

$$A = \left[\hat{d}_1^{t_1}\left(x^{t_1}, y^{t_1}\right) \ \ldots \ \hat{d}_N^{t_1}\left(x^{t_1}, y^{t_1}\right)\right]',$$

$$B = \left[\hat{d}_1^{t_2}\left(x^{t_2}, y^{t_2}\right) \ \ldots \ \hat{d}_N^{t_2}\left(x^{t_2}, y^{t_2}\right)\right]'$$

The values in A and B are bounded from below at unity. To reflect the distance function values about the boundaries in two-dimensional space, Simar and Wilson (1999) form the $(4N \times 2)$ matrix represented in partitioned form by

$$\Delta = \begin{bmatrix} A & B \\ 2 - A & B \\ 2 - A & 2 - B \\ A & 2 - B \end{bmatrix}'$$

The matrix Δ contains 4N pairs of values corresponding to the two time periods. The estimated covariance matrix $\widehat{\Omega}$ below measures the temporal correlation of the original data $[A \ \ B]$ and the estimated covariance matrix of the reflected data $[2 - A \ \ 2 - B]$.

$$\widehat{\Omega} = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_1^2 \end{bmatrix}$$

Moreover the corresponding estimate of the covariance matrix of $[2 - A \ \ B]$ and $[A \ \ 2 - B]$is given by

$$\widehat{\Omega}_R = \begin{bmatrix} \hat{\sigma}_1^2 & -\hat{\sigma}_{12} \\ -\hat{\sigma}_{12} & \hat{\sigma}_1^2 \end{bmatrix}$$

Denote by Δ_j the jth row of Δ, the kernel estimator of the density of the 4N reflected data points represented by the rows of Δ is given by:

$$\hat{g}(z) = \frac{1}{4Nh^2} \sum_{j=1}^{4N} K_j\left(\frac{z - \Delta_j}{h}\right)$$

where $z = [z_1 \ \ z_2]$, and $K_j(.)$ the bivariate normal density function with shape $\widehat{\Omega}$ for $\mathrm{j} = 1, \ldots, \mathrm{N};\ 2\mathrm{N} + 1, \ldots, 3\mathrm{N}$ or shape $\widehat{\Omega}_R$ for $\mathrm{j} = \mathrm{N} + 1, \ldots, 2\mathrm{N};\ 3\mathrm{N} + 1, \ldots, 4\mathrm{N}$ and h is the bandwidth in a kernel density estimator. The smoothing bandwidth parameter (h) was determined by the normal reference rule as suggested by Silverman (1986) for bivariate data

and given by Simar and Wilson (1999) $h = (4/5n)^{\frac{1}{6}}$. Hence, a consistent estimate of the density of the original data $[A \;\; B]$ with bounded support is given by:

$$\hat{g}^*(z) = \begin{cases} 4\hat{g}(z) & for\ z_1 \geq 1,\ \ z_2 \geq 1 \\ 0 & otherwise \end{cases}$$

The second step consist to constructing a pseudo-data set $\tilde{\varphi} = \{(\tilde{x}_{it}, \tilde{y}_{it}); i = 1, \dots n; t = 1,2\}$ to obtain the reference bootstrap technology by using the bivariate kernel density estimation and the reflection method proposed by Silverman (1986) where the bandwidth was selected following the normal reference rule. In fact, to generate the pseudo-samples, Simar and Wilson (1998, 1999) adapt the univariate reflection method described by Silverman (1986) to the bivariate case. This method consist firstly to form the $(N \times 2)$ matrix $\hat{\Delta} = [\delta_{ij}],\ \ i = 1, \dots, N, j = 1,2$ by randomly drawing with replacement from Δ such that each row of Δ has equal probability of selection. The second step consist to compute the $(N \times 2)$ matrix Γ:

$$\Gamma = \frac{1}{\sqrt{(1+h^2)}}\left(\Delta^* + h\rho^* - C\begin{bmatrix} \bar{\delta}_{.1} & 0 \\ 0 & \bar{\delta}_{.2} \end{bmatrix}\right) + C\begin{bmatrix} \bar{\delta}_{.1} & 0 \\ 0 & \bar{\delta}_{.2} \end{bmatrix}$$

where $\bar{\delta}_{.j} = N^{-1}\sum_{i=1}^{N}\delta_{ij}$ for $j = 1,2$ and C is an $(N \times 2)$ matrix of ones, which gives an $(N \times 2)$ matrix of bivariate deviates from the estimated density of Δ, scaled to have the first and second moment properties observed in the original sample represented by $[A \;\; B]$. In addition, ρ^* is an $(N \times 2)$ matrix containing N independent draws from of a normal bivariate density (the kernel functions $K_j(.)$) with the *i*th row of ρ^* representing (i) a draw from a normal density with shape $\hat{\Omega}$ if Δ_i^* was drawn from $[A \;\; B]$ or $[2-A \;\; 2-B]$; or (ii) a draw from a normal density with shape $\hat{\Omega}_R$ if Δ_i^* drawn from $[2-A \;\; B]$ or $[A \;\; 2-B]$.

Draws from a bivariate $N(0, \hat{\Omega})$ density can be simulated by generating independent, identically distributed pseudorandom $N(0,1)$ deviates (z_1, z_2)[3]. The Cholesky decomposition of the (2×2) matrix $\hat{\Omega}$ yields the lower triangular matrix

$$L = \begin{bmatrix} l_1 & 0 \\ l_2 & l_3 \end{bmatrix},$$

[3] using the Box-Muller method (e.g., Press et al.,1986)

Where $LL' = \widehat{\Omega}$, $l_1 = \hat{\sigma}_1$, $l_2 = \hat{\sigma}_{12}/\hat{\sigma}_1$ and $l_3 = (\hat{\sigma}_2^2 - \hat{\sigma}_{12}^2/\hat{\sigma}_1^2)^{\frac{1}{2}}$. Then $(l_1 z_1, l_2 z_2 + l_3 z_3) \sim \mathrm{N}(0, \widehat{\Omega})$. Draws from a $\mathrm{N}(0, \widehat{\Omega}_R)$ density can be simulated similarly by computing $(l_1 z_1, -l_2 z_1 + l_3 z_2) \sim \mathrm{N}(0, \widehat{\Omega}_R)$.

Finally, for each element γ_{ij} of Γ, set

$$\gamma_{ij}^* = \begin{cases} \gamma_{ij}, & if\ \gamma_{ij} \geq 1, \\ 2 - \gamma_{ij} & otherwise \end{cases}$$

The resulting (2×2) matrix $\Gamma^* = [\gamma_{ij}^*]$ consists of two column-vectors of simulated distance function values. Pseudosamples $\tilde{\varphi}$ are then constructed by setting $\tilde{y}_{it_j} = \gamma_{ij}^* y_{ij} / \hat{d}_i^{t_j}\left(x^{t_j}, y^{t_j}\right)$ and $\tilde{x}_{it_j} = x_{it_j}$ for $i = 1, \dots, N, j = 1, 2$

In the third step we calculate the bootstrap estimate of the MPI $\widehat{M}_{ib}^*(t_1, t_2)$ for each firm i through the pseudo-sample attained in step2. Then, repeating steps 2 and 3 for a large number of B times to provide B sets of estimates $\{\widehat{M}_{i1}^*(t_1, t_2), \dots, \widehat{M}_{iB}^*(t_1, t_2)\ \}$ for each firm. Finely, from the bootstrap sample, compute the bias-corrected estimates and confidence intervals for the MPI.

Let $\widehat{M}(t_1, t_2)$ the estimate of the Malmquist index and by $\widehat{M}^*(t_1, t_2)$ the bootstrap estimate of the index and $M(t_1, t_2)$ the *true* unknown index. A starting point to construct confidence intervals of the MPI consist to approximate the unknown distribution of $\widehat{M}(t_1, t_2) - M(t_1, t_2)$ via the distribution of $\widehat{M}^*(t_1, t_2) - \widehat{M}(t_1, t_2)$. Hence, a_α and b_α defining the $(1 - \alpha)$ confidence interval $Pr\left(b_\alpha \leq \widehat{M}(t_1, t_2) - M(t_1, t_2) \leq a_\alpha\right) = 1 - \alpha$. This confidence interval can be approximated by estimating the values a_α^* and b_α^* given by:

$$Pr\left(b_\alpha^* \leq \widehat{M}^*(t_1, t_2) - \widehat{M}(t_1, t_2) \leq a_\alpha^*\right) = 1 - \alpha$$

Thus, an estimated $(1 - \alpha)$ percentage confidence interval for the *i*th MPI is given by:

$$\widehat{M}_i(t_1, t_2) + a_\alpha^* \leq M_i(t_1, t_2) \leq \widehat{M}_i(t_1, t_2) + b_\alpha^*$$

Recall that a value of one for the MPI indicate no productivity change between the two periods t_1 and t_2. Thus, if the interval in (14) does not include unity, the MPI for the *i*th firm is said to be significantly different from unity at α% level. Simar and Wilson (1999) showed that the bootstrap bias estimate for the original estimator $\widehat{M}_i(t_1, t_2)$ is given by:

$$\widehat{bias}_i\left[\widehat{M}_i(t_1, t_2)\right] = B^{-1} \sum_{b=1}^{B} \widehat{M}_{ib}^*(t_1, t_2) - \widehat{M}_i(t_1, t_2)$$

Thus, a bias-corrected estimate of $M_i(t_1, t_2)$ can be computed as:

$$\widetilde{M}_i(t_1, t_2) = 2\widehat{M}_i(t_1, t_2) - B^{-1} \sum_{b=1}^{B} \widehat{M}^*_{ib}(t_1, t_2)$$