Economics Bulletin

Volume 36, Issue 3

Efficiency and productivity measurement with persistent benchmarks

Benjamin Hampf
Darmstadt University of Technology

Abstract

In this note we propose a methodology to estimate the efficiency and productivity of decision making units based on persistent benchmarks. Using panel data we identify convex combinations of observations which dominate the evaluated unit over an extended period of time. Thereby, our approach takes into account that the restructuring of production processes or the adjustment of economic policy according to the optimal benchmark may not be possible instantaneously. We apply our model to an analysis of persistent development benchmarks using a sample of Sub-Saharan countries.

I am grateful to Jens Krüger as well as an anonymous referee for valuable comments. Of course, all remaining errors are mine.
Citation: Benjamin Hampf, (2016) "Efficiency and productivity measurement with persistent benchmarks", *Economics Bulletin*, Volume 36, Issue 3, pages 1715-1721
Contact: Benjamin Hampf - hampf@vwl.tu-darmstadt.de.
Submitted: April 27, 2016. **Published:** September 03, 2016.

1 Introduction

Benchmarking methods have become widely applied tools for micro- and macroeconomic analyses of efficiency and productivity (see Liu et al. (2013) for a survey). However, the results of these analyses crucially depend on the reference observations (or convex combinations thereof) which form the benchmark for the evaluated decision making unit (DMU). Since these benchmarks are used, for example, to derive optimal changes in the production structure of firms or the economic policy of countries, their identification is of major importance for both theoretical and applied research. Traditional approaches like the conventional Data Envelopment Analysis (DEA) have been criticized for using the farthest located point on the technological frontier as the relevant benchmark for a DMU (see e.g. González and Álvarez (2001)) while practitioners refrain from using best-practice benchmarking if the costs associated with reaching the overall frontier are too high (see e.g. Ungan (2004)). In addition, when using panel data to analyze efficiency dynamics and productivity changes traditional methods estimate optimal benchmarks separately for each period. If these benchmarks are very heterogeneous, the policy implications of the analyses may not be consistent over time.

To overcome these drawbacks we propose a model for efficiency and productivity analyses that is based on persistent benchmarks. Using panel data our methodology identifies benchmarks which dominate the DMU under evaluation over an extended period of time. Thereby, our approach is not based on estimating the overall technology frontier in each period but locates benchmarks which provide efficiency improvement potentials for multiple periods. Thus, it allows for time delays when restructuring production processes or implementing new economic policies according to the benchmarks. Based on these reference observations, we calculate quasi-distance functions and quasi-Malmquist indices as proposed by Grifell-Tatjé et al. (1998) to quantify the efficiency and productivity of DMUs.

We apply our model to estimate optimal development benchmarks for Niger, the least developed country in the world, using a sample of 42 Sub-Saharan countries taking into account economic as well as human development indicators. Contrasting the results with those obtained by applying a conventional DEA approach, we find that for the period of 1980-2011 our model identifies a persistent benchmark for Niger which is based on three reference countries while the DEA model leads to vastly varying benchmarks which are constructed by a total of 11 countries.

This note unfolds as follows: Section 2 presents the methodology of our approach to persistent benchmarks while the results of our application are presented and discussed in section 3. Finally, section 4 concludes the note.

2 Methodology

We consider a production process where m inputs $\boldsymbol{x}^t \in \mathbb{R}^m_{++}$ are used to produce k outputs $\boldsymbol{y}^t \in \mathbb{R}^k_{++}$ in each time period $t = 1, \ldots, T$.[1] The technology set of period t comprises all

[1] We assume strictly positive input and output data to ensure that the following programming problems have a feasible solution.

technically feasible input-output combinations in period t and reads as[2]

$$T^t = \left\{ \left(\boldsymbol{x}^t, \boldsymbol{y}^t\right) \in \mathbb{R}_{++}^{m+k} : \boldsymbol{x}^t \text{ can produce } \boldsymbol{y}^t \right\}. \tag{2.1}$$

On this set we impose the conventional axioms for production technologies (e.g. strong disposability of inputs and outputs, convexity) proposed by Shephard (1970) and discussed in Färe and Primont (1995). To construct the technology and to estimate persistent benchmarks we apply nonparametric methods which do not impose a specific functional form of the frontier of the technology set but estimate a piecewise linear envelopment of the empirical observations.

Given a sample of $i = 1, \ldots, n$ DMUs with input-output combinations $(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t)$ the optimal persistent benchmarks can be estimated subject to the nonparametric technology set by solving the nonlinear programming problem

$$\begin{array}{rcll}
\max\limits_{\boldsymbol{s}_x^t, \boldsymbol{s}_y^t, c^t, \boldsymbol{\lambda}} & & \sum_{t=1}^{T} \boldsymbol{1}_m' \frac{\boldsymbol{s}_x^t d_x}{\boldsymbol{x}_i^t} + \boldsymbol{1}_k' \frac{\boldsymbol{s}_y^t d_y}{\boldsymbol{y}_i^t} & \\
s.t.\ \boldsymbol{x}_i^t & = & \boldsymbol{X}^t \boldsymbol{\lambda} c^t + \boldsymbol{s}_x^t & t = 1, \ldots, T \\
\boldsymbol{y}_i^t & = & \boldsymbol{Y}^t \boldsymbol{\lambda} c^t - \boldsymbol{s}_y^t & t = 1, \ldots, T \\
\boldsymbol{1}_n' \boldsymbol{\lambda} & = & 1 & \\
\boldsymbol{s}_x^t, \boldsymbol{s}_y^t, \boldsymbol{\lambda} & \geqq & \boldsymbol{0} & t = 1, \ldots, T \\
c^t & \geqq & 0 & t = 1, \ldots, T.
\end{array} \tag{2.2}$$

Here, $\boldsymbol{X}^t$ $\left(\boldsymbol{Y}^t\right)$ denotes the $m \times n$ $(k \times n)$ matrix of inputs (outputs) of all DMUs in period t. $\boldsymbol{s}_x^t$ $\left(\boldsymbol{s}_y^t\right)$ denotes the $m \times 1$ $(k \times 1)$ vector of slacks. These slacks measure the input excess (output shortage) of the evaluated DMU compared to the benchmark. The benchmark is constructed by a convex combination of DMUs using weights contained in the $n \times 1$ vector $\boldsymbol{\lambda}$ which are restricted to sum up to unity. Note that these weights are independent of the time indices and, hence, the convex combination remains the same for each period. However, by using the scalar c^t we allow for size differences between the evaluated DMU and the persistent benchmark over time. The transposed vectors of ones ($\boldsymbol{1}_m, \boldsymbol{1}_k$ and $\boldsymbol{1}_n$), where subscripts indicate the dimension of the vectors, are used to sum up the slacks and the weight factors. In our model the optimal persistent benchmark is estimated by maximizing the sum of slacks, hence the potential to increase the efficiency of DMU i. In order to render our model unit invariant we follow Charnes et al. (1987) and divide each slack by the actual amount of the corresponding input or output. Due to the non-negativity constraints on the slacks, the benchmark dominates the evaluated DMU in each time period. Nonetheless, the persistent benchmark needs not be located on the overall frontier of the technology set.

The scalars $d_x \in \{0, 1\}$ and $d_y \in \{0, 1\}$ are chosen by the researcher to define the orientation of efficiency measurement. In our following discussion we assume an output orientation which implies $d_x = 0$ and $d_y = 1$.[3] Note that in model (2.2) we assume that the convex combination remains unchanged over the whole sample period. This assumption can be relaxed (e.g. to account for structural breaks) by applying (2.2) to subperiods in the sample.

[2] This is also called the contemporaneous technology set of period t. For alternative specifications of technology sets with panel data see Tulkens and Vanden Eeckaut (1995).

[3] For a discussion of distance functions and Malmquist indices in a slacks-based mixed-oriented model ($d_x = d_y = 1$) see Liu and Wang (2008).

Denoting $\gamma^t = 1/c^t$, $\widetilde{\boldsymbol{s}}_x^t = \boldsymbol{s}_x^t \gamma^t$ and $\widetilde{\boldsymbol{s}}_y^t = \boldsymbol{s}_y^t \gamma^t$ the nonlinear problem (2.2) can be transformed into the linear programming problem

$$\begin{array}{rcll}
\max\limits_{\widetilde{\boldsymbol{s}}_x^t, \widetilde{\boldsymbol{s}}_y^t, \gamma^t, \boldsymbol{\lambda}} & & \sum_{t=1}^{T} \mathbf{1}_m' \frac{\widetilde{\boldsymbol{s}}_x^t d_x}{\boldsymbol{x}_i^t} + \mathbf{1}_k' \frac{\widetilde{\boldsymbol{s}}_y^t d_y}{\boldsymbol{y}_i^t} & \\
s.t.\ \gamma^t \boldsymbol{x}_i^t & = & \boldsymbol{X}^t \boldsymbol{\lambda} + \widetilde{\boldsymbol{s}}_x^t & t = 1, \ldots, T \\
\gamma^t \boldsymbol{y}_i^t & = & \boldsymbol{Y}^t \boldsymbol{\lambda} - \widetilde{\boldsymbol{s}}_y^t & t = 1, \ldots, T \\
\mathbf{1}_n' \boldsymbol{\lambda} & = & 1 & \\
\widetilde{\boldsymbol{s}}_x^t, \widetilde{\boldsymbol{s}}_y^t, \boldsymbol{\lambda} & \geqq & \mathbf{0} & t = 1, \ldots, T \\
\gamma^t & \geqq & 0 & t = 1, \ldots, T.
\end{array} \tag{2.3}$$

The slacks obtained by solving these programming problems can be used to quantify the inefficiency of the DMUs relative to their persistent benchmarks. To aggregate the individual slacks into a scalar measure of inefficiency we follow Grifell-Tatjé et al. (1998) and apply a quasi-distance function which in case of an output-oriented analysis is defined as

$$QD_i^t \left(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t \right) = \left[1 + \mathbf{1}_k' \frac{\boldsymbol{s}_y^t \left(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t \right)}{\boldsymbol{y}_i k} \right]^{-1}. \tag{2.4}$$

Here, $\boldsymbol{s}_y^t \left(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t \right)$ denotes the vector of optimal output slacks for DMU i in period t. A DMU operates output-efficient (inefficient) compared to its persistent benchmark if $QD_i^t \left(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t \right) = 1$ ($QD_i^t \left(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t \right) < 1$).

Based on the quasi-distance function, productivity changes of DMU i between consecutive time periods can be estimated by calculating the quasi-Malmquist index

$$QM_i^{t,t+1} \left(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t, \boldsymbol{x}_i^{t+1}, \boldsymbol{y}_i^{t+1} \right) = \frac{QD_i^t \left(\boldsymbol{x}_i^{t+1}, \boldsymbol{y}_i^{t+1} \right)}{QD_i^t \left(\boldsymbol{x}_i^t, \boldsymbol{y}_i^t \right)}. \tag{2.5}$$

Here, $QD_i^t \left(\boldsymbol{x}_i^{t+1}, \boldsymbol{y}_i^{t+1} \right)$ denotes the mixed-period quasi-distance function which indicates the inefficiency of the input-output combination of period $t+1$ relative to the persistent benchmark in period t. It can be obtained by solving the programming problem

$$\begin{array}{rcll}
\max\limits_{\boldsymbol{s}_x^t, \boldsymbol{s}_y^t, \boldsymbol{s}_x^{t+1}, \boldsymbol{s}_y^{t+1}, c^t, \boldsymbol{\lambda}} & & \sum_{t=1}^{T} \left(\mathbf{1}_m' \frac{\boldsymbol{s}_x^t d_x}{\boldsymbol{x}_i^t} + \mathbf{1}_k' \frac{\boldsymbol{s}_y^t d_y}{\boldsymbol{y}_i^t} \right) + \sum_{t=1}^{T-1} \left(\mathbf{1}_m' \frac{\boldsymbol{s}_x^{t,t+1} d_x}{\boldsymbol{x}_i^{t+1}} + \mathbf{1}_k' \frac{\boldsymbol{s}_y^{t,t+1} d_y}{\boldsymbol{y}_i^{t+1}} \right) & \\
s.t.\ \boldsymbol{x}_i^t & = & \boldsymbol{X}^t \boldsymbol{\lambda} c^t + \boldsymbol{s}_x^t & t = 1, \ldots, T \\
\boldsymbol{x}_i^{t+1} & = & \boldsymbol{X}^t \boldsymbol{\lambda} c^t + \boldsymbol{s}_x^{t,t+1} & t = 1, \ldots, T-1 \\
\boldsymbol{y}_i^t & = & \boldsymbol{Y}^t \boldsymbol{\lambda} c^t - \boldsymbol{s}_y^t & t = 1, \ldots, T \\
\boldsymbol{y}_i^{t+1} & = & \boldsymbol{Y}^t \boldsymbol{\lambda} c^t - \boldsymbol{s}_y^{t,t+1} & t = 1, \ldots, T-1 \\
\mathbf{1}_n' \boldsymbol{\lambda} & = & 1 & \\
\boldsymbol{s}_x^t, \boldsymbol{s}_y^t, \boldsymbol{\lambda} & \geqq & \mathbf{0} & t = 1, \ldots, T \\
c^t & \geqq & 0 & t = 1, \ldots, T \\
\boldsymbol{s}_x^{t,t+1}, \boldsymbol{s}_y^{t,t+1} & & \text{free} & t = 1, \ldots, T-1.
\end{array} \tag{2.6}$$

Here, $\boldsymbol{s}_x^{t+1}$ and $\boldsymbol{s}_y^{t+1}$ represent the mixed-period slacks of DMU i. Note that since the input-output combination of period $t+1$ may be located outside the technology of period t these slacks may be negative. The quasi-Malmquist index exhibits a value larger than 1 (lower than 1) if the productivity of DMU i has increased (decreased) between two time periods.

3 Empirical application

We apply our model to an analysis of optimal development benchmarks for Sub-Saharan countries. In our presentation we focus on the results for the least developed country Niger. Nonparametric methods are particularly suited for this evaluation since they allow to include multiple inputs *and* outputs (see Mariano et al. (2015) for an overview on development-related efficiency studies). Therefore, we may not only account for economic but also human development without the necessity to introduce a fixed weighting scheme as done in the Human Development Index (HDI).[4]

In our analysis we include two inputs (capital stock and labor force) as well as the outputs GDP as a measure of economic development and the years of primary and secondary schooling as well as the life expectancy at birth as measures of human development. Data on inputs and the GDP have been obtained from the Penn World Tables 8 while data on the indicators of human development have been obtained from the World Development Indicators database provided by the Worldbank. Our dataset includes 42 Sub-Saharan countries and covers the period from 1980 to 2011. Descriptive statistics of the data are presented in table I. A list of the included countries as well as the ISO-3 codes can be found in table II in the appendix.

Table I: Descriptive statistics ($n = 42$ Sub-Saharan countries)

Variable	Unit	Min	Mean	Median	Max	St.dev.
Capital stock	Mio. 2005 $	249.33	43240.65	14754.87	952435.15	100074.78
Labor force	Mio. worker	0.03	4.87	2.62	50.63	7.23
GDP	Mio. 2005 $	121.20	20815.17	7734.03	392611.09	47079.81
Schooling	Years	9.00	12.39	12.00	15.00	0.72
Life expectancy	Years	27.08	52.80	52.30	73.27	7.16

The results of our analysis are depicted in figure 1. The upper panel presents the optimal λ-values to construct the benchmarks for Niger. The lines (solid, dashed and dotted) represent the weights as obtained from our approach which identifies persistent benchmarks while the weights from a conventional output-oriented DEA analysis are indicated by points. The optimal development benchmark for Niger based on our model is a convex combination of Comoros (COM), Gabon (GAB) and Zimbabwe (ZWE). Note that these observations are not necessarily located on the overall frontier. The optimality principle of our model only requires that the convex combination of these observations dominates the country under evaluation, in this case Niger. By construction the persistent benchmark weights are constant over time. In contrast, the results for the conventional approach show that the benchmarks vary largely over time. 11 countries are used to construct the benchmark for at least one year. While some of them are parts of the benchmark for a short period and afterwards disappear (e.g. Togo), others vary in terms of their relevance for the convex combination (e.g. Mauritius). Hence, in contrast to our model this analysis does not indicate consistent reference points to orient at in order to increase the performance of Niger in terms of economic and human development.

[4] Using optimal, unrestricted weights for development variables, as is done in nonparametric approaches, is also critically discussed in the literature. See Bougnol et al. (2010) for an efficiency analysis with development indicators including weight restrictions.

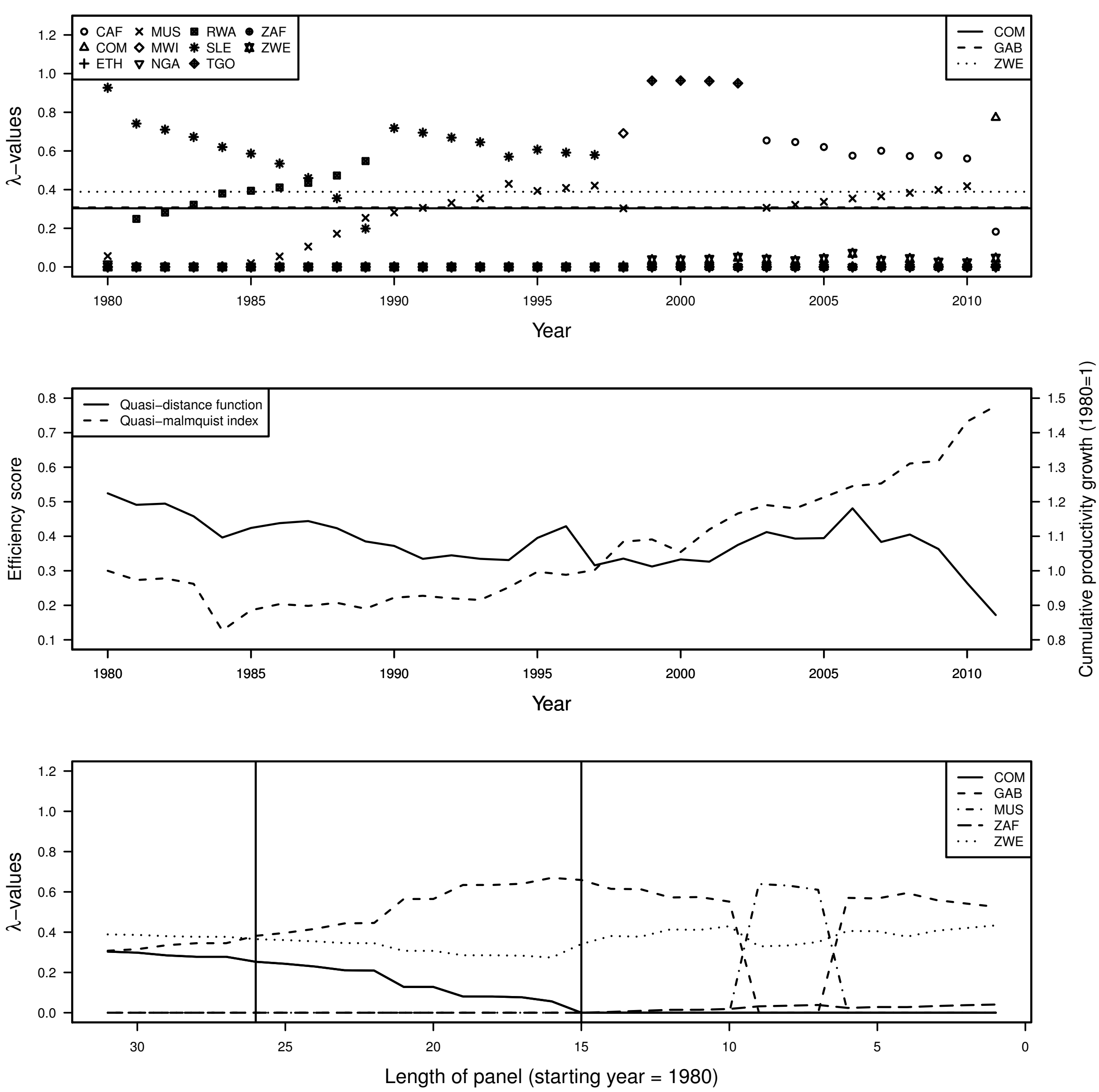

Figure 1: Results for Niger (1980-2011)

The middle panel presents the efficiency and productivity results for Niger based on the persistent benchmark. The solid line represents the results for the quasi-distance function while the dashed line represents the cumulative productivity results based on the quasi-Malmquist index. The efficiency of Niger shows a declining trend. Nonetheless, compared to 1980 Niger has increased its productivity by 47% until 2011. This implies that Niger was able to enhance its human and economic development over time. However, the persistent benchmark increased its level of development even further leading to a larger inefficiency. This result is in line with the findings for the HDI (see UNDP (2015)) which show that although Niger was able to increase its HDI results over time it still remains the least developed country.

To address the sensitivity of the persistent benchmark selection with regard to the length of the analyzed panel, the lower graph presents the persistent benchmark results for varying numbers of included years. If all 31 years of observations from the sample are included in the analysis, the benchmark corresponds to the persistent benchmark presented in the upper graph. Decreasing the number of analyzed time periods influences the persistent benchmark, but the effect is far less pronounced than the heterogeneity for the conventional DEA analysis presented above. For example, the left vertical line in the figure indicates that up to 5 years can be removed from the analysis without changing the reference DMUs which form the persistent benchmark or their ranking in terms of the λ-factors. Moreover, as indicated by the right vertical line, up to 16 years can be removed before the reference DMUs change. Hence, our model provides very homogeneous reference points to evaluate the economic and development performance of Niger to, even if the length of the analyzed panel is significantly reduced.

4 Conclusion

In this note we have developed a new methodology to measure efficiency and productivity based on persistent benchmarks. We applied the approach to identify persistent benchmarks among Sub-Saharan countries for the economic and human development of Niger. Future research may extend the model to decompose the productivity changes into efficiency and technical change along the lines of Färe et al. (1994). Moreover, in our application we focus on identifying persistent benchmarks for Niger based on economic and human development performance. Future research may extend the analysis by accounting for external factors (e.g. quality of institutions, geographic location) following the approach by Golany and Thore (1997). Finally, since in contrast to previous approaches our model leads to stable benchmarks, the general idea of forecasting Malmquist indices and productivity changes as proposed by Daskovska et al. (2010) could be modified to take this additional information into account in order to increase the precision of the forecasts.

References

Bougnol, M.-L., J. H. Dulá, M. P. Estellita Lins, and A. C. Moreira da Silva (2010). "Enhancing Standard Performance Practices with DEA". In: *Omega* 38, pp. 33–45.

Charnes, A., W. W. Cooper, J. Rousseau, and J. Semple (1987). "Data Envelopment Analysis and Axiomatic Notions of Efficiency and Reference Sets". Research Report CCS 558, Center for Cybernetic Studies, University of Texas, Austin.

Daskovska, A., L. Simar, and S. Van Bellegem (2010). "Forecasting the Malmquist Productivity Index". In: *Journal of Productivity Analysis* 33, pp. 97–107.

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

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang (1994). "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries". In: *American Economic Review* 84, pp. 66–83.

Golany, B. and S. Thore (1997). "Restricted Best Practice Selection in DEA: An Overview with a Case Study Evaluating the Socio-Economic Performance of Nations". In: *Annals of Operations Research* 73, pp. 117–140.

González, E. and A. Álvarez (2001). “From Efficiency Measurement to Efficiency Improvement: The Choice of a Relevant Benchmark”. In: *European Journal of Operational Research* 133, pp. 512–520.

Grifell-Tatjé, E., C. A. K. Lovell, and J. T. Pastor (1998). “A Quasi-Malmquist Productivity Index”. In: *Journal of Productivity Analysis* 10, pp. 7–20.

Liu, F.-H. F. and P.-H. Wang (2008). “DEA Malmquist Productivity Measure: Taiwanese Semiconductor Companies”. In: *International Journal of Production Economics* 112, pp. 367–379.

Liu, J. S., L. Y. Y. Lu, W.-M. Lu, and B. J. Y. Lin (2013). “A Survey of DEA Applications”. In: *Omega* 41, pp. 893–902.

Mariano, E. B., V. A. Sobreiro, and D. A. do Nascimento Rebelatto (2015). “Human Development and Data Envelopment Analysis: A Structured Literature Review”. In: *Omega* 54, pp. 33–49.

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

Tulkens, H. and P. Vanden Eeckaut (1995). “Non-Parametric Efficiency, Progress and Regress Measures for Panel Data: Methodological Aspects”. In: *European Journal of Operational Research* 80, pp. 474–499.

UNDP (2015). *Human Development Report 2015*. United Nations Development Programme.

Ungan, M. (2004). “Factors Affecting the Adoption of Manufacturing Best Practices”. In: *Benchmarking: An International Journal* 11, pp. 504–520.

Appendix

Table II: List of countries

ISO-3	**Country**	**ISO-3**	**Country**	**ISO-3**	**Country**
AGO	Angola	GIN	Guinea	NER	Niger
BDI	Burundi	GMB	Gambia	NGA	Nigeria
BEN	Benin	GNB	Guinea-Bissau	RWA	Rwanda
BFA	Burkina Faso	GNQ	Equatorial Guinea	SEN	Senegal
BWA	Botswana	KEN	Kenya	SLE	Sierra Leone
CAF	Central African Republic	LBR	Liberia	STP	São Tomé and Principe
CIV	Côte d'Ivoire	LSO	Lesotho	SWZ	Swaziland
CMR	Cameroon	MDG	Madagascar	TCD	Chad
COG	Congo	MLI	Mali	TGO	Togo
COM	Comoros	MOZ	Mozambique	TZA	Tanzania
CPV	Cabo Verde	MRT	Mauritania	UGA	Uganda
ETH	Ethiopia	MUS	Mauritius	ZAF	South Africa
GAB	Gabon	MWI	Malawi	ZMB	Zambia
GHA	Ghana	NAM	Namibia	ZWE	Zimbabwe