



Technology-based total factor productivity and benchmarking: New proposals and an application [☆]

Mircea Epure ^a, Kristiaan Kerstens ^{*,b}, Diego Prior ^{a,b}

^a Department of Business Economics, Universitat Autònoma de Barcelona, Spain

^b CNRS-LEM (UMR 8179), IESEG School of Management 3 rue de la Digue, F-59800 Lille, France

ARTICLE INFO

Article history:

Received 7 March 2010

Accepted 6 January 2011

Processed by B. Lev

Available online 19 January 2011

Keywords:

DEA

Banking

ABSTRACT

The present study fills a gap between the benchmarking literature and multi-output based efficiency and productivity studies by proposing a benchmarking framework to analyze total factor productivity (*TFP*). Different specifications of the Hicks–Moorsteen *TFP* index are tailored for specific benchmarking perspectives: (1) static, (2) fixed base and unit, and (3) dynamic *TFP* change. These approaches assume fixed units and/or base technologies as benchmarks. In contrast to most technology-based productivity indices, the standard Hicks–Moorsteen index always leads to feasible results. Through these specifications, managers can assess different facets of the firm's strategic choices in comparison with firm-specific relevant benchmarks and thus have a broad background for decision making. An empirical application for the Spanish banking industry between 1998 and 2006 illustrates the managerial implications of the proposed framework.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Literature on benchmarking focuses on the selection of a unit of strategic value against which performance is compared [1]. Another series of academic studies analyze the efficiency and productivity of firms with multiple inputs and outputs. So far there seems little or no link between these two streams of research. In this paper we propose to bridge this gap by defining novel total factor productivity (*TFP*) benchmarking methods. These are devised to include cross-sectional and inter-temporal perspectives not only concerning unit to unit benchmarking, but also efficiency frontier benchmarking. These various perspectives are introduced stepwise starting with static indices, continuing with fixed base and unit, and ending with dynamic benchmarking. This provides managers of any industry with a new set of *TFP* benchmarking indices for decision making.

Both benchmarking and *TFP* analysis represent key tools in business economics. For instance, Balk [2] points to two main actions a manager constantly carries out: the monitoring activity (i.e., assessing how the firm is doing over time) and the benchmarking

activity (i.e., comparing firm performance with respect to its main competitors). Although both activities aim at enhancing performance, monitoring is internally oriented while benchmarking has an external focus.

Benchmarking is defined as the search and emulation of the industry's best practices and it thus is an objective setting procedure [1]. Through benchmarking, a firm can deduce whether it has a best or worst practice. Thus, it can aim at maintaining superiority or at closing the gap to its competitors [1]. Therefore, benchmarking appeals most to firms with similar strategic orientations or facing comparable problems and opportunities [3,4].

Empirical applications suggest different methods for monitoring or benchmarking activities. In managerial studies of performance, the simplest method is the use of output-input ratios or any other kind of ratios for that matter (see [5,6]). Managers care about profitability and implicitly about productivity: “the most encompassing measure of productivity change, *TFP* change, is nothing but the “real” component of profitability change. Put otherwise, if there is no effect of prices then productivity change would coincide with profitability change” ([2]: 6).

The above *TFP* measures are easily adaptable to benchmarking purposes. One can simply divide the firm's *TFP* change (or performance) ratio to the one of a chosen competitor. However, in multiple inputs and outputs technologies various problems emerge related to the use of ratios for benchmarking. When comparing two firms, different partial productivity ratios (built by dividing different outputs by some inputs) can point to different results. The management literature suggests a way to remedy this problem.

[☆]We thank two referees, the editor-in-chief and conference participants at EWEP 2009 and IVCEP-EFIUCO 2009 for most constructive comments that substantially improved the paper. The usual disclaimer applies. The authors acknowledge financial support from the Spanish Ministerio de Ciencia e Innovación (Grant ECO2010-18967).

* Corresponding author.

E-mail address: k.kerstens@ieseg.fr (K. Kerstens).

Specifically, in the presence of prices, multiple outputs and inputs productivity indices are proposed by the American Productivity Center (APC) method [7].

Turning attention to efficiency and productivity analysis, this literature uses frontier methods with economic underpinning in production theory to handle multiple inputs yielding multiple outputs. These non-parametric techniques have known an important upsurge and are probably best known under the label Data Envelopment Analysis (DEA) (see [8,9]). DEA methods compute the degree of inefficiency separating a certain Decision Making Unit (DMU) from the efficiency frontier. In this case, the comparison is done against the whole analyzed sample, not against some specific strategic competitor as in benchmarking. Thus, in DEA benchmarks are the efficient units on the frontier against which the other DMUs are projected using some efficiency measure (see [8,9]). Therefore, it is highly unlikely that a single benchmark is found for all units evaluated in the sample.

In inter-temporal analyses, the efficiency and productivity literature captures the potentially shifting efficiency frontier usually through index numbers. The Malmquist productivity index is probably the best known measure that has been extensively used in past research.¹ However, there are some pitfalls to the use of Malmquist indices. First, it is not always a TFP index: while the TFP properties are maintained under constant returns to scale, shortcomings appear in the presence of variable returns to scale (VRS) which mostly represents the true technology [13]. Second, there is the possibility of having infeasible results.² For example, Glass and McKillop [15] find infeasibilities for up to 7% of the analyzed UK building societies.³ This issue could have an important impact on benchmarking analysis, since managers wish to obtain firm level results that may not always be available.⁴

As a result, there are two main issues with the Malmquist index that need to be resolved: TFP interpretation and infeasibilities. To address these problems, one can turn to Bjurek's [19] proposal for a Hicks–Moorsteen TFP (HMTFP) index (see also [20]: footnote 18). The HMTFP index is defined as a ratio of an aggregate output quantity over an aggregate input quantity index. More precisely it measures the change in output quantities in the output direction and the change in input quantities in the input direction, instead of exclusively adopting an input- or output-orientation as Malmquist indices usually do. The TFP characteristics of the HMTFP index solve the limitations of the traditional Malmquist productivity index in the presence of VRS. Furthermore, this HMTFP index is well-defined under general assumptions of variable returns to scale and strong disposability.⁵

¹ See the general survey of Färe et al. [10], the survey on the banking sector in Fethi and Pasiouras [11], or applications/decompositions as the one of Wheelock and Wilson [12].

² The literature sometimes gives the impression that imposing constant returns to scale eliminates the issue of infeasibility. However, Briec and Kerstens [14] demonstrate that constant returns to scale are a necessary, but not a sufficient condition to guarantee that the Malmquist index is well-defined.

³ Yörük and Zaim [16] report infeasible computations that reach 10% of their sample of OECD countries. Also, for the Spanish insurance industry Cummins and Rubio-Misas [17] mention that infeasibilities are present (without indicating the exact amount).

⁴ To solve the problem of infeasibilities, Kao [18] propose a common-weights global Malmquist productivity index: apart from the common weights (i.e., the same frontier facet for every DMU), this amounts to creating a common frontier for all DMUs in all time periods.

⁵ Briec and Kerstens [21] demonstrate that the Hicks–Moorsteen productivity index satisfies the determinateness property under mild conditions. According to Bjurek ([19]: 310) the feasibility of this index is attributable to the property that “all input efficiency measures included meet the condition that the period of the technology is equal to the period of the observed output quantities” and “all output efficiency measures included meet the condition that the period of the technology is equal to the period of the observed input quantities”.

However, in spite of its attractive properties, the HMTFP has been scarcely empirically applied.⁶

Various benchmarking applications have been developed in the non-parametric efficiency and productivity analysis framework by isolating reference frontiers or DMUs. In the non-TFP context, Berg et al. [25] adapt the Malmquist productivity index to have a base year frontier as a benchmark frontier, and measure productivity growth or regress relative to this fixed basis. Similarly Berg et al. [26] adapt the Malmquist productivity index to make comparisons across countries with respect to a fixed basis (i.e., a single country) for a given year. Also, single benchmark TFP analyses have been undertaken by Zaim et al. [27], Färe et al. [28] and Zaim [29]. Manipulating a Hicks–Moorsteen index, their proposals include both cross-sectional and inter-temporal analyses by mixing a single DMU and TFP benchmarking. Zaim et al. [27] use a five years sample of OECD countries to analyze the well-being of individuals in each country as compared to a benchmark country. Similarly environmental performance is measured against a benchmark DMU in Färe et al. [28] and Zaim [29]. While the former study looks upon OECD countries at cross-sectional level, the latter analyzes US states from both cross-sectional and inter-temporal perspectives.

A small existing literature thus proposes efficiency frontier comparisons using productivity indices combined with some form of unit to unit benchmarking. But, while consensus is reached regarding the usefulness of benchmarking, less agreement exists with respect to the choice of benchmarks. In a strategic analysis setting, the interest of a firm may be to know its relative performance to a certain specific competitor, instead of comparing itself to a frontier potentially composed of all firms in the sector. The benchmark could differ for each firm, even though it could remain the same over a certain time period. In addition, awareness of TFP positioning is useful in both static and dynamic environments. Efficiency coefficients (static) and TFP indices (dynamic) relative to a given benchmark are equally relevant and could represent the basis of strategic decision making. For instance, in the case of similar strategic configurations, firms constitute strategic groups and may choose their benchmark within their relevant cluster. In this case, the benchmark unit can be the leader of the strategic group or any other unit, say the local competitor, regardless of its performance.

To develop a systematic framework to analyze these issues, this study proposes a TFP benchmarking framework by adapting Bjurek's [19] HMTFP index for benchmarking purposes. The introduced HMTFP indices for benchmarking include the features of the traditional HMTFP together with some of the properties of the indices in Berg et al. [25,26], Zaim et al. [27], Färe et al. [28], and Zaim [29]. Various specifications of the HMTFP index measure distances (and catching-up effects) between analyzed DMUs and their selected benchmarks: these indices offer TFP interpretations with respect to static, fixed base or changing efficiency frontiers.

The empirical application considers the Spanish banking sector over the period 1998–2006, a post-deregulation growth phase. The sector experienced consistent growth following the disappearance of regulatory constraints and due to the competition between private and savings banks. In productivity and efficiency terms, the sector has been looked at from a multitude of perspectives.⁷

⁶ Bjurek et al. [22] is the first empirical application of the Hicks–Moorsteen index. To the best of our knowledge, there are only two more empirical applications/decompositions of the Hicks–Moorsteen index: one is developed in a parametric context by Nemoto and Goto [23], another is proposed in O'Donnell [24].

⁷ E.g., Grifell-Tatjé and Lovell [30,31], Lozano-Vivas [32], Kumbhakar and Lozano-Vivas [33], Más-Ruiz et al. [34], Tortosa-Ausina et al. [35], or Illueca et al. [36].

In addition, there is a wide range of *DEA* studies that focus on some alternative aspects of benchmarking. We mention some recent examples. Bournol et al. [37] show how *DEA* can be used by practitioners to enhance standard performance evaluations such as benchmarking or constructing rankings based on scorecard assessments. Moreover, the versatility of *DEA* models for benchmarking allows to evaluate multiple-stakeholder perspectives using common sets of variables [38]. In a similar vein, *DEA*-based benchmarking can also be used for analyzing bank branch efficiency suitable for both line managers and senior executives [39]. However, our study is unique in focusing on integrating a benchmarking perspective into frontier-based *TFP* measures.

This paper is structured as follows. Section 2 develops the *HMTFP* index adapted to benchmarking purposes. Section 3 presents sample information together with the variables and methods of analysis. The empirical application is found in Section 4, while the final section is dedicated to some concluding remarks.

2. The Hicks–Moorsteen *TFP* index adapted to benchmarking

2.1. The Hicks–Moorsteen *TFP* index and its interpretation

Caves et al. [40] introduced the Malmquist index into the mainstream literature as a ratio of either output or input distance functions. This index is based on technology information only (i.e., output and input quantities) and requires no price information. Furthermore, this index is always partially oriented (either output or input). Following some cursory remarks in the earlier literature (see [20]), Bjurek [19] introduces the technology-based Hicks–Moorsteen productivity index that combines output and input quantity indices defined using output and input distance functions, respectively, making it simultaneously oriented.

For period t , let us define an input vector $x_t \in R_+^N$ and an output vector $y_t \in R_+^M$ forming the technology T_t of feasible input–output combinations. The input distance function in period t is defined as

$$D_t^i(y_t, x_t) = \max_{\theta} \{ \theta > 0 : (x_t/\theta, y_t) \in T_t \} \tag{1}$$

The input distance function “treats (multiple) outputs as given, and contracts input vectors as much as possible consistent with the technological feasibility of the contracted input vector” ([8]: 10). This function presents a complete description of the structure of multi-input, multi-output efficient production technology. Furthermore, it offers “a complete characterization of the structure of multi-input, multi-output efficient production technology, and it provides a reciprocal measure of the distance from each producer to that efficient technology” ([8]: 10).

The output distance function in period t can be defined as

$$D_t^o(y_t, x_t) = \min_{\phi} \{ \phi > 0 : (x_t, y_t/\phi) \in T_t \} \tag{2}$$

This output distance function has similar characteristics, and can be equally employed to characterize the structure of efficient production technologies in the multi-output case [8]. These distance functions can be defined using general specifications of technology (e.g., a non-parametric technology with variable returns to scale).

The basic *HMTFP* index [19] based on a technology in year t and computing changes between observations in periods $t(y_t, x_t)$ and $t+1(y_{t+1}, x_{t+1})$ is defined as follows:

$$HMTFP_t = \frac{D_t^o(y_{t+1}, x_t)/D_t^o(y_t, x_t)}{D_t^i(y_t, x_{t+1})/D_t^i(y_t, x_t)} \tag{3}$$

In line with Bjurek’s [19] proposal, the above distance functions are evaluated with respect to a technology assuming *VRS* and strong disposability of inputs and outputs. The *HMTFP* index shows the shifts in the technology between two analyzed periods,

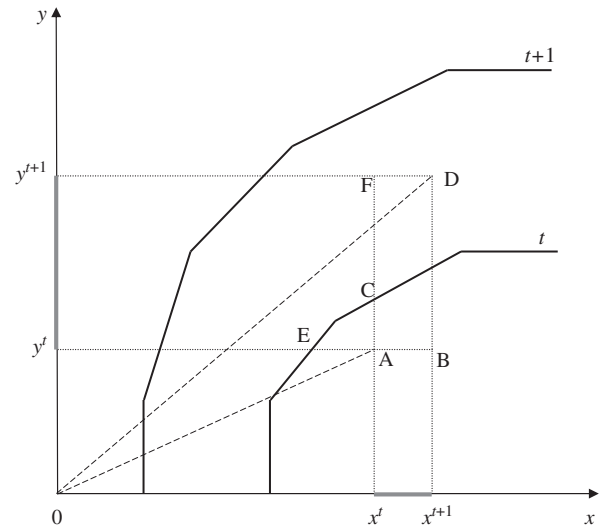


Fig. 1. The Hicks–Moorsteen total factor productivity index (adapted from Bjurek et al. ([22]: 223)).

both compared against the technology in the first year. The *HMTFP* scores are to be read in line with other ratio-based indices: specifically, values greater than one indicate *TFP* growth, whereas values lower than one point to decreases in *TFP*.

In the one input one output case, productivity is equal to the division of a single output over a single input (y/x), whereas productivity change is the quotient of two productivity ratios—in $t+1$ and t —($(y_{t+1}/x_{t+1})/(y_t/x_t)$) (see [2,7]). In the multiple inputs and outputs case, a *TFP* index is required to obtain a similar interpretation for a general technology (see Grifell-Tatjé and Lovell [13], Balk [2], and O’Donnell [24]). The advantage of a *TFP* index is that it provides information on the movements in productivity by comparing multi-dimensional real output growth and input growth (see [2]). Thus, the *HMTFP* is among the frontier-based index numbers having a correct *TFP* interpretation, since it divides a multidimensional index of real output growth by an index of real input growth.

Fig. 1 illustrates the *HMTFP* index in expression (3) with respect to the technology at time t . Assume that point A represents a *DMU* at time t and point D the same *DMU* at time $t+1$. One can compute this *DMU*’s *TFP* change by following the specification of the output and input distance functions in (1) and (2), respectively, jointly with the *HMTFP* index definition in (3). The numerator of the *HMTFP*, the output quantity index, is the ratio of the vertical line segments x^tF/x^tC to x^tA/x^tC . The denominator of the *HMTFP*, the input quantity index, is the ratio of the horizontal line segments y^tB/y^tE to y^tA/y^tE . Furthermore, in line with the *TFP* interpretation of the index, the same result can be attained by dividing the values of the two productivity ratios $((y_{t+1}/x_{t+1})/(y_t/x_t))$. It is worth pointing out that both methods are describing the changes in outputs and inputs to explain the transition from points A to D. Finally this result is nothing else than the division of the two slopes corresponding to points A and D (see dotted lines in Fig. 1).⁸

While the *HMTFP* index coincides with the traditional *TFP* interpretation, it also reveals one aspect that ratios are not able to show: information referring to the efficiency frontier. This is shown by the numerator and the denominator of the *HMTFP* index. First, the output quantity index provides in the output direction an efficient frontier benchmark (point C, as indicated by

⁸ See Appendix 1 (electronic supplementary material) for a detailed illustration of the *HMTFP* index and its properties by means of a numerical example.

segment x^C). At the same time, it is shown how the *DMU* is not efficient in period t ($D_t^o(y_t, x_t) = x^A/x^C < 1$). Second, in the input direction, one finds point E as the benchmark (as indicated by segment y^E) and also the input inefficiency in period t ($D_t^i(y_t, x_t) = y^A/y^E > 1$). These movements on the output or input side represent the distances needed to reach a specific point on the best practice frontier.

All the above measurements are done with respect to the technology in the first year. Establishing a one year technology instead of selecting a geometric mean index is common practice in the benchmarking literature (see the Malmquist index in Berg et al. [25] or the *HMTFP* index in Zaim et al. [27], Färe et al. [28], and Zaim [29]). For instance, Berg et al. [25] make use of a base technology to obtain a fixed benchmark for measuring technical change. This fixed base Malmquist index has the added advantage that it is transitive (circular), while its geometric mean version is not.⁹ Similarly, Zaim et al. [27] propose an improvement index defined through a one year technology and using as benchmark a *DMU* in the analyzed time period. In line with the above authors, the reason not to combine technologies is quite straightforward: when performing benchmarking analysis the benchmark should be well determined and easy to identify.

There is one more aspect worthwhile mentioning. Criticism can be targeted to the pseudo-observations created by the *HMTFP* index, some of whose components are defined by including different time periods in the same distance function. This can be observed in expression (3) where outputs in periods $t+1$ or t are combined with inputs in periods t or $t+1$, respectively. Nevertheless, these mixed time periods are a main characteristic of the *HMTFP* index and contribute to both its *TFP* interpretation as well as to its feasibility. These combinations can further appear in benchmarking adaptations in the form of distance functions containing outputs (inputs) from one *DMU* and inputs (outputs) from another.

2.2. Adapting the *HMTFP* index to benchmarking purposes: three proposals

It is now important to clearly delimitate possible benchmarking approaches. While the introductory section explains the motives for choosing a single unit as a benchmark, there are still pros and cons for each possible specification. The adaptations of the *HMTFP* index for benchmarking compare the productivity of two different *DMUs* in a variety of contexts. First, a static index provides a distance between analyzed *DMUs* and their benchmark. Second, the comparison is done against a fixed *DMU* and a base technology frontier. This is useful for situations in which managers achieve a good understanding of a competitor in a certain time period, and by iterating computations over the years they can observe the eventual catching-up effects that have been attained. Third, the dynamic benchmarking perspective is developed by contrasting *TFP* changes between analyzed *DMUs* and their benchmarks while allowing for both to evolve over time. The latter definition is novel in the efficiency benchmarking literature and helpful to capture catching-up effects which account for changes in technology.

2.2.1. The static *HMTFP* index for benchmarking

The static adaptation of the *HMTFP* index for benchmarking can be mathematically expressed as follows:

$$HMTFP_{st} = \frac{D_t^o(y_t, x_t^B)/D_t^o(y_t^B, x_t^B)}{D_t^i(y_t^B, x_t)/D_t^i(y_t, x_t^B)} \quad (4)$$

⁹ As noticed in footnote 4 of Berg et al. [25] the far more popular geometric mean specification of the Malmquist index does not have much support in the original Caves et al. [40] paper.

where t is the only period under analysis, (y_t, x_t) are the outputs and inputs of the analyzed *DMU* in period t , and (y_t^B, x_t^B) are the outputs and inputs of the unit established as a benchmark. This specification of the *HMTFP* index permits one to compute, for a certain period t , the distance from each *DMU* to an established benchmark point (B). A similar approach with a fixed base unit has been defined for the Malmquist index in Berg et al. [26].

*HMTFP*_{st} index values higher than unity indicate that the analyzed *DMU* has a higher *TFP* than its benchmark, whereas values lower than unity point out a worse performer. In this way, the scores quantify the *TFP* advantage a *DMU* has with respect to its benchmark in a certain period or the catching-up it needs to reach this reference point.¹⁰

2.2.2. The fixed base *HMTFP* index for benchmarking

The above static *HMTFP* index for benchmarking (see similar applications in Färe et al. [28] or Zaim [29]) has, however, one pitfall: it does not include a time component. Traditionally this problem was solved by defining a base year (benchmark technology) dynamic index (see, e.g., the fixed base Malmquist index in Berg et al. [25]). By combining the fixed base index with the single *DMU* benchmarking, the fixed base and unit *HMTFP* is specified as

$$HMTFP_{fbk} = \frac{D_k^o(y_t, x_k^B)/D_k^o(y_k^B, x_k^B)}{D_k^i(y_k^B, x_t)/D_k^i(y_k^B, x_k^B)} \quad (5)$$

where k is the (constant) base year and t is the year under analysis, (y_t, x_t) are the outputs and inputs of the analyzed *DMU* in period t , and (y_k^B, x_k^B) are the outputs and inputs of the *DMU* established as benchmark (fixed in the base year).

In contrast to the static case, it is now possible to see movements over time with respect to the *DMU* set as benchmark. Both the technology frontier and the benchmark are kept fixed in period k . Therefore, by computing changes between period k and period $t, t+1$, etc. one is examining shifts in the technology with respect to a known position set as a goal for the evaluated *DMU*. Therefore, the *HMTFP*_{fb} may show higher or lower than unity results. A higher/lower than unity score indicates the percentage in which a *DMU* performs better/worse in terms of *TFP* in the analyzed period, as compared to its benchmark in the base period.¹¹

The advantage of this second option is the availability of *TFP* changes over time with respect to a benchmark in a base period. However, one could argue against the relevance of fixing the technology at a certain point. Since the technology, the evaluated *DMUs* and the benchmark all change over time, the comparison of a *DMU* with regard to a benchmark in a given base year becomes somewhat obsolete after being used for various periods of time. It is as if one keeps aiming at a target that has meanwhile almost certainly moved onwards.

2.2.3. The dynamic *HMTFP* index for benchmarking: decomposing the *HMTFP*

A third proposal starts from the standard *HMTFP* index (see (3)) and introduces a new decomposition that offers a dynamic viewpoint for benchmarking purposes. This decomposition proposal represents a novelty to the existing literature. The chosen course of action is to decompose the basic *HMTFP* index (3) such that its components are suitable for a dynamic benchmarking analysis.

¹⁰ See Appendix 1 (electronic supplementary material) for a numerical illustration of the static *HMTFP* index for benchmarking.

¹¹ See Appendix 1 (electronic supplementary material) for a numerical illustration of the fixed base *HMTFP* index for benchmarking.

Through simple mathematical rearrangement, the *HMTFP* index in expression (3) can be decomposed as follows:

$$\begin{aligned}
 HMTFP_t &= \frac{D_t^o(y_{t+1}, x_t) / D_t^o(y_t, x_t)}{D_t^i(y_t, x_{t+1}) / D_t^i(y_t, x_t)} \\
 &= \frac{(D_t^o(y_{t+1}, x_t) / D_t^o(y_t, x_t)) / (D_t^o(y_{t+1}^B, x_t^B) / D_t^o(y_t^B, x_t^B))}{(D_t^i(y_t, x_{t+1}) / D_t^i(y_t, x_t)) / (D_t^i(y_{t+1}^B, x_{t+1}^B) / D_t^i(y_t^B, x_t^B))} \\
 &\quad \underbrace{\hspace{10em}}_{TFP \text{ change relative to the benchmark}} \\
 &\quad \times \underbrace{\frac{D_t^o(y_{t+1}^B, x_{t+1}^B) / D_t^o(y_t^B, x_t^B)}{D_t^i(y_t^B, x_{t+1}^B) / D_t^i(y_t^B, x_t^B)}}_{\text{Benchmark's } TFP \text{ change}} \quad (6)
 \end{aligned}$$

where t and $t+1$ are the years under analysis, (y_t, x_t) and (y_{t+1}, x_{t+1}) are the outputs and inputs of the analyzed *DMU* in periods t and $t+1$, respectively, and (y_t^B, x_t^B) and (y_{t+1}^B, x_{t+1}^B) are the outputs and inputs of the unit established as a benchmark in periods t and $t+1$, respectively.¹²

This dynamic *HMTFP* index and its two decomposition elements offer a two-fold benchmarking perspective. The first decomposition component is called the *TFP* change relative to the benchmark. This is simply the division of the *HMTFP* of the analyzed *DMU* to the benchmark's *HMTFP*. This ratio of *HMTFP* indices relates the *TFP* change of the analyzed *DMU* to the *TFP* change of the benchmark. Thus, it basically computes the variation in *TFP* between the *DMU* under analysis and the benchmark unit. Consequently values higher (lower) than unity reflect that the *TFP* of the analyzed *DMU* changes faster (slower) than the one of the benchmark.

The second decomposition component is simply the benchmark's *HMTFP* index. For our purpose, it is called the benchmark's *TFP* change (as indicated below the braces in (6)), and it simply indicates how the *TFP* of the benchmark varies over time.

Note again that the *HMTFP* index on the *LHS* shows the *TFP* changes between t and $t+1$ (see (3) and its interpretation) and that there is no benchmark involved at all. The advantage of this decomposition approach is that it combines both the frontier-based *TFP* analysis and the benchmarking approach. Furthermore, by running this decomposition over several consecutive time periods, statistical tests between its components may reveal catching-up or falling-behind effects for the first and second components relative to the frontier and the benchmark, respectively.

2.2.4. Synthesis

Thus, each of the three adaptations of the *HMTFP* index (expressions (4), (5), and (6)) offers a certain benchmarking scenario. Naturally a manager or regulator can select the most appropriate method for his/her specific situation and needs. While each of these three approaches can stand alone, these methods are also potentially complementary. In the latter case, a multidimensional perspective can be obtained via the parallel interpretations of these three *HMTFP* indices for benchmarking.

3. Sample description and specification issues

3.1. Description of the Spanish banking industry

The Spanish banking industry proves to be attractive for research due to its rapid growth and global competition between different bank types. This growth occurred after the second half of the 1980s, triggered by the deregulation of the sector [41]. The

year 1989 marks the start of the liberalized Spanish banking market where those earlier viewed as small intermediaries could now act in ways similar to private banks. The savings banks benefited most from these policies, since apart from the permission to perform general banking operations they were allowed to expand throughout Spain. Consequently it was probably the savings banks' strategic choice of expansion that led to the global competition between private and savings banks still manifest today.

However, the years 1992–1996 represented a difficult time for this sector. In 1995, towards the end of this period, a key step for this industry's development was taken through the introduction of a novel legal regime for bank creation. The sector introduced novel technologies (e.g., important increases of *ATMs*' networks, information systems) together with the establishment of new financial products and services [41]. Moreover, at the end of the 1990s the annual reports of the savings banks reveal a clear strategic choice for expansion, mostly through opening new branches. For instance, Illueca et al. [36] find productivity gains related to savings banks that expand outside their original markets. This strong option for growth implies the adaptation of the management of inputs and outputs to new forms of organization. Accordingly the end of the 1990s represents a cornerstone for growth and is attractive to analyze when developing new lines of research.

Considering the ownership composition, there are three types of banking institutions in Spain: private banks, savings banks, and credit cooperatives. The market belongs with a vast majority to the first two categories, while only a fraction of the banking activity remains in the hands of the credit cooperatives. The private banks, which are shareholder-oriented, generally pursue the goal of profit maximization. By contrast, the savings banks are organized as foundations and they include boards of trustees formed by representatives of regional authorities, city halls, employees, depositors and the founding entity. Finally the credit cooperatives frequently belong to their customers. It is important to emphasize two main differences between the credit cooperatives and the other two bank types. First, there are important size dissimilarities, since credit cooperatives are a lot smaller. Second and partly related, technology is homogenous between the private and savings banks only. The credit cooperatives are less developed not only in terms of branch (geographical) reach, but also in terms of *ATMs* and other products and services.

The above discussion yields two conclusions. First, the analysis' starting point is the year 1998. This corresponds to the end of the financial difficulties in a deregulated Spanish banking sector. It also stands for the beginning of a novel growth period defined by new corporate strategies, particularly in the case of savings banks. Second, the homogeneity of the employed technology is guaranteed by forming a sample of private and savings banks (and excluding the credit cooperatives).

3.2. Specification issues: input and output variables

Banking studies provide various ways to define the outputs and inputs for productivity and efficiency analyses. Studies reviewing the input and output variables employed in banking are those of Berger and Humphrey [42], Goddard et al. [43], or Fethi and Pasiouras [11]). There are two main approaches to the choice of how to measure the flow of services provided by financial institutions: these are the production and the intermediation approaches. The production approach considers banks as producers of deposit and loan services. When considering this specification, just physical inputs like labor and capital and their costs must be included. In contrast, the intermediation approach

¹² See Appendix 1 (electronic supplementary material) for a numerical illustration of the dynamic *HMTFP* index for benchmarking.

Table 1
Summary of outputs–inputs median levels, 1998–2006.

| Year | No. of units | Loans y_1 | Securities y_2 | Non-interest income y_3 | Deposits x_1 | Operating assets x_2 | Labor (no. of employees) x_3 | Other operating expenses x_4 |
|-------|--------------|-------------|------------------|---------------------------|----------------|------------------------|--------------------------------|--------------------------------|
| 1998 | 73 | 1749,396 | 172,346 | 23,722 | 2327,570 | 191,458 | 1035 | 25,182 |
| 1999 | 73 | 1985,205 | 206,504 | 27,951 | 2472,112 | 231,313 | 1104 | 24,024 |
| 2000 | 73 | 2454,182 | 254,494 | 32,179 | 2877,276 | 235,096 | 1171 | 24,846 |
| 2001 | 73 | 2710,152 | 298,812 | 33,564 | 3207,999 | 305,994 | 1202 | 26,416 |
| 2002 | 73 | 3175,310 | 341,328 | 35,726 | 3699,197 | 276,829 | 1232 | 28,237 |
| 2003 | 73 | 3847,798 | 345,866 | 37,437 | 4219,151 | 301,926 | 1248 | 29,937 |
| 2004 | 73 | 4372,738 | 451,945 | 43,437 | 4668,497 | 335,834 | 1251 | 32,524 |
| 2005 | 73 | 5768,165 | 689,706 | 45,724 | 5378,141 | 340,653 | 1334 | 34,680 |
| 2006 | 73 | 6980,295 | 775,457 | 53,264 | 6612,898 | 351,860 | 1409 | 36,887 |
| Total | 657 | 3293,806 | 315,664 | 36,895 | 3588,932 | 286,962 | 1198 | 28,088 |

All values excepting x_3 in thousands of Euros in constant prices. x_3 in absolute numbers.

regards banks as intermediaries through which deposits and purchased funds are transformed into loans and financial investments. Hence, under this framework funds and their interest cost (which are the raw material to be transformed) have to be introduced as inputs in the model.

This analysis uses the intermediation approach, indicated by the current banking and efficiency studies as the best way to describe bank activity (see Berger and Humphrey [42] or Goddard et al. [43]). Moreover, recent studies of the Spanish banking industry employ various types of intermediation approaches to measure the flow of services (see e.g., [41,45,36]). In this paper, the employed definition of variables and the reasoning behind its choice are similar to the ones of Illueca et al. [36]. However, it is not identical due to the fact that Spanish accounting formats changed between 2004 and 2005, and the resulting need to maintain homogeneity for all of the analyzed periods. The selected outputs are: (1) loans, (2) securities and (3) non-interest income (non-traditional output). Inputs are (1) deposits, (2) operating assets, (3) labor (number of employees), and (4) other operating expenses.

With the exception of labor, all variables are in monetary terms (thousands of Euros). The reason for this design is quite straightforward. Let us consider two banks having an equal number of deposits, although the monetary quantity in one bank is twice as in the other. In this case, accounting for the deposits in monetary terms is more relevant for showing which bank holds a larger output. For these monetary values, the numbers are deflated with respect to the GDP. Ultimately labor is used in absolute numbers, as these values prove higher consistency throughout the analyzed sample and produce less bias.

Prior to setting up the final sample, a test for outliers has been performed. It is well-known that extreme points can influence the shape of the estimated production frontier and introduce bias in the *TFP* changes. In relation to frontier analysis, two influential contributions are those of Andersen and Petersen's [44] super-efficiency coefficient and Wilson's [45] paper. Therefore, through the super-efficiency test the influential units found in the sample are removed and the efficiency measures re-estimated. Moreover, following Prior and Surroca [46], this procedure is continued as long as the null hypotheses of equality between successive efficiency scores cannot be rejected. By doing so, about 6% of the banks in the total sample turn out to be outliers and are therefore eliminated.

In addition, foreign banks that show inconsistent assets-related information have also been removed. Considering data availability, the calculations are performed on a yearly basis between 1998 and 2006. By balancing the panel corresponding to this period 1998–2006, a final sample of 73 private and savings banks is formed. Table 1 offers descriptive statistics for each year from 1998 to 2006 as well as for the overall period. Notice the substantial growth of both median output and input levels over time.

Before computing the different specifications of the *HMTFP* indices for benchmarking, one peer *DMU* must be selected. As previously mentioned, this choice should be in accordance with each bank's strategic options and competitive positioning. For the following application, two up to date best practice criteria are selected: (1) technical efficiency and (2) good performance in the Committee of European Bank Supervisors (*CEBS*) stress test published in July 2010 by the Bank of Spain.¹³ The bank that fully complies with these criteria is Bancaja, a medium-sized savings bank. First, Bancaja is technically efficient during all the analyzed periods. Second, it shows a good capital adequacy ratio in the *CEBS* stress test results.¹⁴ While the minimum acceptable capital adequacy ratio was set to 6% (50% more than the legally required minimum), Bancaja's group obtained a 8.6% ratio.

Moreover, Bancaja represents an interesting benchmark since during the analyzed period it experienced important evolutions from a management viewpoint. First, it was involved in a merger. Second, it consistently followed a strategy to expand outside its region. These two last aspects are illustrated in more detail when analyzing the empirical results.

The differences between the successive *HMTFP* indices are assessed through a Li test (see [47,48]). This non-parametric statistical test compares two unknown distributions using kernel densities. Its advantages are two-fold. First, the Li test statistic is valid for dependent and as well as for independent variables (see [48]). Second, in contrast to most statistical tests, the Li test is not based on mean or median comparisons, but instead compares two entire distributions to each other. Thus, by means of the Li test *p*-value, the null hypothesis of equality of distributions can be rejected or not.

4. Empirical application

This section progressively presents the empirical results of the benchmarking adaptations of the *HMTFP* index. All computed results are feasible and *TFP* interpretations are offered together with frontier components. Specifically one can see the *TFP*

¹³ One can argue that one should avoid setting a very large bank as the benchmark (e.g., Banco Santander), since it could be experiencing diminishing returns to scale (i.e., be positioned in the upper right corner of a graph similar to the one in Fig. 1). In such a configuration, in a static environment it is very likely for the slopes of most *DMUs* (situated in the increasing returns to scale area) to be higher than the one for this very large benchmark. Moreover, a medium-sized *DMU* is probably more suitable as a benchmark, since mimicking its strategy should be easier for most of the analyzed *DMUs*.

¹⁴ As a key part of the *CEBS* stress test, the capital adequacy ratio is calculated to ensure that banks are solvent: they must have sufficient capital to resist under adverse and unlikely conditions. This capital adequacy ratio was computed as the total Tier 1 capital (i.e., core capital, which includes equity capital and disclosed reserves) divided by the total risk-weighted assets.

behavior of the analyzed *DMU* in the *HMTFP* index result and also obtain efficiency frontier information both for output and input orientations from the numerator and denominator of the index. When used for managerial decision making, these features of the *HMTFP* index significantly improve upon the properties of a standard Malmquist index.

4.1. The static *HMTFP* index for benchmarking

The first step of the analysis reports the results of the static *HMTFP_{st}* index for benchmarking (see (4)). Table 2 presents the descriptive statistics associated with the *HMTFP_{st}*, while Fig. 2 shows the comparison between the analyzed sample's median and the benchmark. The results of the *HMTFP_{st}* index are straightforward and easy to understand. Considering that the benchmark's score is equal to 1 by definition in this first case, all the other values show the distance in terms of *TFP* to this benchmark *DMU*. Keep in mind that results higher/lower than unity indicate a better/worse *TFP* than the benchmark.

Table 2 reports the *HMTFP_{st}* index results at sample level. The index's distribution is shown through percentile results, which have the advantage of avoiding the biases that top/bottom *DMUs* can create in the mean values. It is first found that between 1998 and 2001 the *TFP* at the median level is rather similar to the benchmark. However, during the same period the *TFP* differences between the sample and Bancaja have decreased at the 10th and 25th percentiles. Moreover, the 75th percentile shows that the top 25% of the sample increased the static *TFP* distances to Bancaja from 36% in 1998 to 53% in 2001.

In 2002, results at the 75th and 90th percentiles show a smaller positive distance to Bancaja than in the previous year. Also looking at the 10th and 25th percentiles, it is rather clear that the *TFP* measures for the sample were closer to Bancaja in 2002 and 2003. During the following years, the *TFP* of the analyzed *DMUs* relative to Bancaja deteriorates at the 10th,

Table 2
Static *HMTFP* for benchmarking—descriptive statistics.

| Year | <i>HMTFP_{st}</i> percentiles | | | | | <i>HMTFP_{st}</i> mean | <i>HMTFP_{st}</i> std. dev. | Bench's <i>HMTFP_{st}</i> |
|------|---------------------------------------|------|--------|------|------|--------------------------------|-------------------------------------|-----------------------------------|
| | 10 | 25 | Median | 75 | 90 | | | |
| 1998 | 0.71 | 0.84 | 0.98 | 1.36 | 1.72 | 1.13 | 0.41 | 1.00 |
| 1999 | 0.74 | 0.85 | 1.00 | 1.30 | 1.54 | 1.11 | 0.39 | 1.00 |
| 2000 | 0.75 | 0.86 | 1.00 | 1.43 | 1.83 | 1.19 | 0.48 | 1.00 |
| 2001 | 0.80 | 0.89 | 1.02 | 1.53 | 2.37 | 1.35 | 0.77 | 1.00 |
| 2002 | 0.85 | 0.98 | 1.20 | 1.44 | 2.03 | 1.31 | 0.50 | 1.00 |
| 2003 | 0.87 | 0.94 | 1.12 | 1.34 | 1.83 | 1.26 | 0.55 | 1.00 |
| 2004 | 0.77 | 0.85 | 0.99 | 1.21 | 1.56 | 1.17 | 0.64 | 1.00 |
| 2005 | 0.73 | 0.84 | 0.94 | 1.23 | 1.93 | 1.25 | 0.97 | 1.00 |
| 2006 | 0.70 | 0.78 | 0.94 | 1.22 | 2.26 | 1.28 | 1.12 | 1.00 |

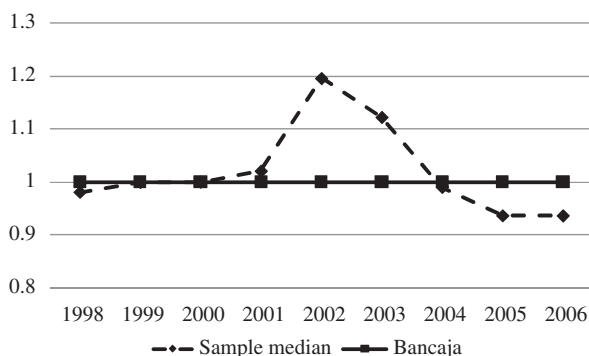


Fig. 2. Static *HMTFP* benchmarking—benchmark Bancaja vs. sample median.

25th, and 50th percentiles, the distances regarding the 75th percentile remains stable, while the 90th percentile experiences an increase in *TFP* relative to Bancaja.

Fig. 2 illustrates the evolution of the median levels of the *HMTFP_{st}* index. The sample median is clearly similar to Bancaja's *HMTFP_{st}* index until 2001 and superior in 2002 and 2003. The Li test confirms that in 2002 and 2003 the *HMTFP_{st}* index is significantly superior to the rest of the years.¹⁵ However, the catching-up of Bancaja starts in 2004 and becomes obvious in 2005 and 2006. There are several reasons that can explain these movements.

On the one hand, in 2001 Bancaja was involved in a merger with a smaller savings bank (Caixa Carlet). Although in monetary terms the dimension of the merger is certainly limited, one can expect some adjustment costs (e.g., creating common procedures and streamlining branch networks may take several years). Such initial decrease of efficiency following mergers and acquisitions has been previously documented for Spanish savings banks by Cuesta and Orea [41]. Similar to our results for Bancaja, this same study also indicates that significant efficiency increases occur a few years after the merger. Alternative explanations for decreases in efficiency are available. For instance, it is well established [49] that mergers and acquisitions can generate “big bath” charges in accounting terms, which represent significant non-recurring losses or operating expenses taken in the current period to open the door for improved future earnings. In so doing, taking advantage of specific events, managers can have an interest in under-reporting earnings in the current period to be capable to report higher performance in the future.

On the other hand, in 2003 Bancaja enhanced its already important expansion process by opening branches in other regions, also a fruitful strategy for increasing efficiency as shown by Illueca et al. [36]. This increased its size by three times in just 6 years. Obviously this output expansion requires an initial effort to invest in installing an additional, initially underutilized capacity. This is exactly what Fig. 2 and Table 2 illustrate for the years 2002 and 2003. Regarding the years 2004–2006, it is evident that this expansion allowed Bancaja to improve its *TFP* far better than the sample median. Notice, however, that these catching-up effects are provided from a static *TFP* benchmarking perspective. For scrutinizing dynamic *TFP* movements, one should employ one of the following two proposals.

4.2. The fixed base *HMTFP* index for benchmarking

Having the yearly snapshots of the Spanish banking sector in mind, the fixed base benchmark analysis is conducted. The *HMTFP_{fb}* index for benchmarking (see (5)) is computed by establishing a base-year technology (i.e., 1998) and fixing the benchmark Bancaja in the same period. Thus, while the benchmark and reference technology are not allowed to change, the sample can experience technological change. In general, this index must be interpreted in the same way as its static version. Nonetheless, there are a few differences. It is now important to compare the scores of the analyzed *DMUs* with the ones of the benchmark shown in the last column of Table 3, since the benchmark's index is no longer unity throughout. This column reveals a constant and very substantial increase in *TFP* of Bancaja over the years relative to the base year, with a small increase in 2002 and even a minor decrease in *TFP* in 2003 (consistent with the information provided in Table 3 for the *HMTFP_{st}* index).

¹⁵ An Appendix 2 (electronic supplementary material) containing detailed significant differences between the analyzed years and a selection of graphical displays of the Li tests is available.

Table 3
Fixed Base $HMTFP_b$ for benchmarking—descriptive statistics.

| Year | $HMTFP_b$ percentiles | | | | | $HMTFP_b$ mean | $HMTFP_b$ std. dev. | Bench's $HMTFP_b$ |
|-------|-----------------------|------|--------|------|------|----------------|---------------------|-------------------|
| | 10 | 25 | Median | 75 | 90 | | | |
| 98–99 | 0.78 | 0.90 | 1.08 | 1.44 | 1.84 | 1.22 | 0.51 | 1.13 |
| 98–00 | 0.90 | 1.01 | 1.15 | 1.61 | 2.18 | 1.36 | 0.53 | 1.29 |
| 98–01 | 0.92 | 1.08 | 1.24 | 1.61 | 2.33 | 1.46 | 0.60 | 1.41 |
| 98–02 | 0.97 | 1.14 | 1.41 | 1.72 | 2.18 | 1.51 | 0.50 | 1.45 |
| 98–03 | 1.05 | 1.23 | 1.47 | 1.85 | 2.58 | 1.63 | 0.60 | 1.44 |
| 98–04 | 1.10 | 1.35 | 1.63 | 1.93 | 2.44 | 1.75 | 0.66 | 1.78 |
| 98–05 | 1.37 | 1.62 | 2.13 | 2.79 | 4.26 | 2.56 | 1.57 | 2.43 |
| 98–06 | 1.56 | 1.90 | 2.42 | 3.01 | 4.38 | 2.73 | 1.39 | 3.10 |

Analyzing the $HMTFP_b$ index at the sample level, the results in Table 3 simply indicate the TFP increase/decrease with respect to the base year technology and benchmark. Thus, one can note that at the median level (as well as any of the percentiles) the Spanish private and savings banks experience a substantial and progressive TFP growth over the period compared to Bancaja in 1998. This improvement of the TFP indicates the quick and generalized positive evolution of the efficiency in the whole industry.

Previous research indicates that the main drivers of productivity and efficiency in Spanish banking are deregulatory measures and technological change. Both factors were mainly linked to savings banks. For instance, Grifell-Tatjé and Lovell [30] found that productivity declined at the end of the 1980s and the beginning of the 1990s, the initial phase of the sector's deregulation. However, Grifell-Tatjé and Lovell [31]—using a different output specification—found productivity growth. Next, Kumbhakar and Lozano-Vivas [33] illustrate that in the long-run (i.e., 1986–2000) this domestic deregulation has had a significant positive effect on savings banks' TFP growth, whereas the European deregulation mainly benefited the private banks' TFP . Moreover, for the post-deregulation period up to 2004, Tortosa-Ausina et al. [35] and Illueca et al. [36] find productivity increases mostly due to technological change. These results are in line with the TFP growth shown in Table 3. We also confirm strong TFP growth till at least the end of our sample (i.e., 2006).

Next, one should compare these sample level TFP results in the first part of Table 3 with the TFP results of Bancaja reported in the last column. To find TFP growth results over the whole period close to Bancaja's, one must be situated around the 75th percentile of the distribution or above. Thus, the evolution of Bancaja's TFP is parallel to roughly the best 25% of the sample. However, the $DMUs$ in the 90th percentile clearly beat the TFP growth track record of Bancaja in a substantial way.

In Fig. 3, one can trace the TFP results at the sample median over the period and compare these to the one of Bancaja. For example, between 1998 and 2002, the sample result at the median level is 1.41 compared to the technology and benchmark in 1998. Given that the benchmark's index is 1.45, there is a difference of 4% ($1.45 - 1.41 = 0.04$) in favor of the benchmark. This indicates that the benchmark realizes a superior TFP growth compared to approximately 50% of the analyzed sample. By comparing the position between both lines in Fig. 3 or by computing these differences, one observes that the sample median is closest to Bancaja in the years 2002 and 2003. Furthermore, as in the case of the $HMTFP_{st}$ index in Section 4.1, this gap between the sample median and Bancaja increases near the end of the analyzed time period. These findings are consistent with the above information justifying Bancaja's movements due to its important expansion plan.

The advantage of this $HMTFP_b$ index in terms of informative content is that bank managers may establish as their goal a

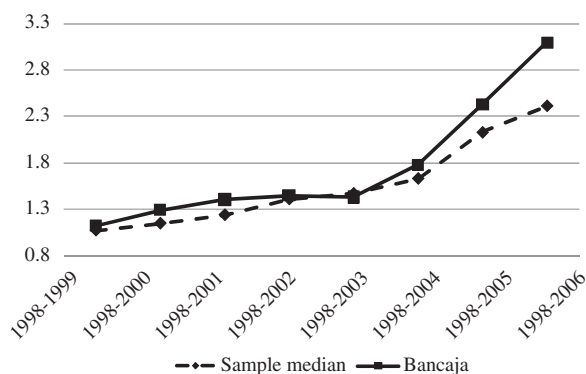


Fig. 3. Fixed base $HMTFP_b$ benchmarking—benchmark Bancaja vs. sample median.

Table 4
Standard $HMTFP$ —descriptive statistics.

| Year | $HMTFP_{dyn}$ (standard) percentiles | | | | | $HMTFP_{dyn}$ mean | $HMTFP_{dyn}$ std. dev. |
|-------|--------------------------------------|------|--------|------|------|--------------------|-------------------------|
| | 10 | 25 | Median | 75 | 90 | | |
| 98–99 | 1.01 | 1.06 | 1.11 | 1.19 | 1.31 | 1.13 | 0.13 |
| 99–00 | 1.01 | 1.09 | 1.15 | 1.20 | 1.38 | 1.18 | 0.18 |
| 00–01 | 0.95 | 1.01 | 1.06 | 1.14 | 1.29 | 1.11 | 0.19 |
| 01–02 | 1.05 | 1.09 | 1.17 | 1.28 | 1.48 | 1.23 | 0.27 |
| 02–03 | 0.95 | 1.05 | 1.10 | 1.19 | 1.27 | 1.15 | 0.21 |
| 03–04 | 1.01 | 1.07 | 1.13 | 1.21 | 1.32 | 1.15 | 0.13 |
| 04–05 | 1.10 | 1.19 | 1.32 | 1.64 | 2.16 | 1.51 | 0.57 |
| 05–06 | 1.06 | 1.12 | 1.16 | 1.22 | 1.31 | 1.18 | 0.11 |

Table 5
Dynamic $HMTFP$ decomposition components—descriptive statistics.

| Year | $HMTFP_{dyn}$ to benchmark percentiles | | | | | $HMTFP_{dyn}$ to bench mean | $HMTFP_{dyn}$ to bench std. dev. | Bench's $HMTFP_{dyn}$ |
|-------|--|------|--------|------|------|-----------------------------|----------------------------------|-----------------------|
| | 10 | 25 | Median | 75 | 90 | | | |
| 98–99 | 0.90 | 0.94 | 0.98 | 1.05 | 1.16 | 1.01 | 0.12 | 1.13 |
| 99–00 | 0.86 | 0.93 | 0.98 | 1.02 | 1.17 | 1.01 | 0.15 | 1.17 |
| 00–01 | 0.86 | 0.90 | 0.95 | 1.02 | 1.15 | 1.00 | 0.17 | 1.12 |
| 01–02 | 0.97 | 1.00 | 1.08 | 1.18 | 1.37 | 1.13 | 0.25 | 1.08 |
| 02–03 | 0.93 | 1.03 | 1.08 | 1.16 | 1.24 | 1.12 | 0.20 | 1.02 |
| 03–04 | 0.85 | 0.91 | 0.96 | 1.03 | 1.12 | 0.97 | 0.11 | 1.18 |
| 04–05 | 0.77 | 0.83 | 0.92 | 1.14 | 1.51 | 1.05 | 0.40 | 1.44 |
| 05–06 | 0.74 | 0.79 | 0.81 | 0.86 | 0.92 | 0.83 | 0.08 | 1.42 |

certain DMU in a certain period and try to reach its position as part of the bank's strategy. However, when pursuing this over longer periods, this does not allow including more recent information about the benchmark. Therefore, such strategy may become little informative after a while, unless one updates either the benchmark and/or the fixed base year. For instance, one may decide to substitute Bancaja by another benchmark starting from a new base year, or one may stick to Bancaja but update the base year to ensure one catches up with the moving target. Alternatively one can try to account for these movements by introducing the following dynamic analysis.

4.3. The $HMTFP$ index and its new decomposition for benchmarking

To document the relevance of our proposal, we illustrate the empirical results for the new $HMTFP_{dyn}$ index in the greatest detail. First, since the $HMTFP_{dyn}$ index starts off from the standard $HMTFP$ index (3), we first report these results in Table 4. Next, we report the new decomposition proposal constituting our $HMTFP_{dyn}$ for benchmarking index in Table 5.

Starting with Table 4, note that no benchmark is involved and results should be interpreted as traditional TFP changes.

Therefore, these results illustrate the yearly evolution of the Spanish banking industry between 1998 and 2006. This evolution is in line with the positive TFP changes shown by the $HMTFP_{st}$ index. It also confirms previous empirical research showing consistent TFP growth for the Spanish banks (e.g., Grifell-Tatjé and Lovell [31], Kumbhakar and Lozano-Vivas [33], Tortosa-Ausina et al. [35], or Illueca et al. [36]).

The percentile levels in Table 4 indicate that a drop in the TFP change only occurs below the 10th percentile or slightly above the 10th percentile and only for years 2001 and 2003. Also at the 10th percentile, the most spectacular growth is 10% in 2005. This same year also shows the most important TFP improvements for the rest of the percentile levels. Regarding the 2005 results, it must be stated that these higher scores may be partly caused by changes introduced in the Spanish accounting formats (see Section 3.2) that create difficulties to maintain the homogeneity for specific variables.

While all TFP changes are positive above the 25th percentile, one can see in Table 4 that fluctuations appear with respect to their amount. Nevertheless, the Li tests find that the TFP growth shown during the last two periods is significantly superior to the one obtained during all other periods.¹⁶ It is also seen that TFP changes between 2000 and 2001 are significantly lower than in the rest of the periods. This may be due to the expansion phase of the Spanish banking sector, since most banks (particularly the savings banks) enhanced their branch networks until the beginning of the 2000s. It is thus probable that banks improved their managerial practices at the end of this process. Note that this implication is analogous to the previous findings for the benchmark (Bancaja) in the case of the $HMTFP_{st}$ and $HMTFP_{fb}$ indices.

Next, Table 5 presents the decomposition results. All descriptive statistics are illustrating the TFP change relative to the benchmark (the first decomposition component in (6)), while the last column of Table 5 presents the second decomposition component, the benchmark's TFP change. This second component is interpreted straightforwardly as the standard $HMTFP_{dyn}$ index of Bancaja, showing positive TFP changes for all periods. However, these changes are lower in 2002 and 2003 and quite higher in 2005 and 2006, all with respect to the other analyzed periods. These results are consistent with the managerial interpretations in Section 4.1.

Nonetheless, the new and appealing component is the TFP change relative to the benchmark. Again, the comparison against a fixed benchmark is the key issue, as this component measures the changes in the TFP of each analyzed DMU relative to the changes in the TFP of the benchmark. Percentiles show that for the first three years, a DMU must be close to the 75th percentile to have TFP changes comparable to Bancaja. It is interesting to notice that for changes between 1999 and 2000, the distance separating the median from the 75th percentile is minimal.

The sample's TFP change relative to the benchmark are quite different during the following two periods. At the end of 2002, it is sufficient to be positioned at the 25th percentile to have TFP changes similar to Bancaja. Moreover, in the same year the 75th percentile indicates a TFP which is 18% above the benchmark, whereas at the 90th percentile this relative gap moves up to 37%. These values are somewhat lower for 2003. Therefore, it is once more revealed how the two years in which the sample is performing better than Bancaja are 2002 and 2003. Thus, they include the effects of the merger event and the expansion process, as suggested in Section 4.1.

¹⁶ Even if some mean or median levels are similar, significant differences appear as the Li test is based on dissimilarities between the entire distributions. Table A2.3 in Appendix 2 (electronic supplementary material) illustrates these differences.

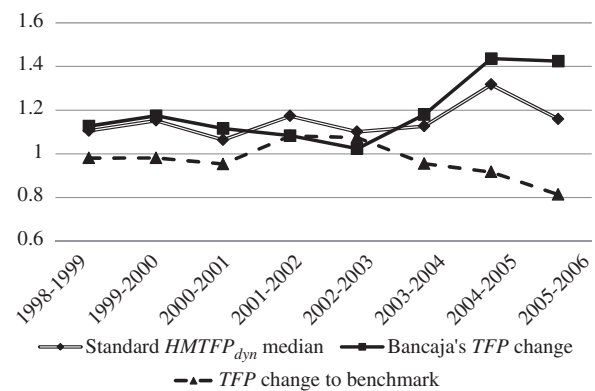


Fig. 4. Dynamic $HMTFP$ decomposition—benchmark Bancaja vs. sample median.

Bancaja's relative TFP growth is consistent throughout 2004–2006 at the 10th, 25th and 50th percentiles. At median level, Bancaja is superior by 4% in 2004, but this gap increases to 19% in 2006. However, these last periods offer a few surprising results. First, in 2006 there is a difference of only 2% between the 25th percentile and the median. Second, for 2004 Bancaja's TFP growth is outpacing TFP growth of the $DMUs$ situated in the 75th and 90th percentiles. For instance, in 2006 Bancaja's superiority at the 90th percentile amounts to 8%.

Looking at the benchmark's TFP change in the last column of Table 5, we observe that Bancaja experiences high TFP growth in the beginning of the period. A substantial slow down appears in the years 2002 and 2003, again confirming earlier results. Thereafter, Bancaja seems to be accelerating its TFP growth, which only weakens in 2006.

This dynamic picture of the Spanish banking sector can also be observed in Fig. 4, which presents the decomposition results of the $HMTFP_{dyn}$ index. When read jointly, these results lead to competitive advantage interpretations. First, the standard $HMTFP_{dyn}$ index median levels represent the fluctuations in TFP growth, which is highest in 2005. Second, the TFP change relative to the benchmark median levels are represented by the dashed line. Notice how this dashed line is above unity in 2002 and 2003 and ends up at almost 20% below this level in 2006. Yet again it is graphically shown how Bancaja's catching-up materializes starting with 2004.¹⁷ Lastly the solid line shows that Bancaja has its lowest TFP growth in 2003, with spectacular increases in 2005, later maintained for 2006.

Fig. 4 makes apparent that there are differences between a sole DMU 's analysis (as the benchmark's TFP change) and the analysis of sample level results. The next subsection demonstrates the importance of these unit level analyses.

4.4. Unit to unit analysis

After progressively advancing into the TFP -benchmarking analysis of the Spanish banking sector, the proposed methodology can present a global picture by combining the three approaches. As already indicated, these independent TFP benchmarking indices can be combined to achieve complementary perspectives for managers and regulators. While sample-level results are probably most relevant for regulators and researchers, bank managers are arguably more interested in comparing a particular bank to the benchmarking unit. For this second type of analysis and in the presence of VRS , the feasibility of $HMTFP$ indices offers a clear advantage over the standard Malmquist indices. To illustrate this

¹⁷ Using Appendix 2 (electronic supplementary material) one can confirm that most of the observed differences are significant.

unit to unit approach, we maintain Bancaja as the benchmark, and follow two specific DMUs that partially comply with the benchmark criteria in Section 3.2. We select two medium-sized banks that are technically efficient in most of the analyzed periods and with good stress test results: Banca March (abbreviated *BM*), a private bank with a capital adequacy ratio of 19.7% (best ratio in the 2010 Bank of Spain test), and Caja de Guipuzkoa y San Sebastian (abbreviated *GSS*), a savings bank with a capital adequacy ratio of 13%.

Figs. 5–8 show the three proposed benchmarking approaches applied to these two specific DMUs. These figures can be interpreted similarly to the previous ones, with the only difference that, instead of median levels, one is now evaluating single units relative to the benchmark. Thus, this analysis takes a DMU-specific viewpoint as it presents all results combined first for *BM* and then for *GSS*.

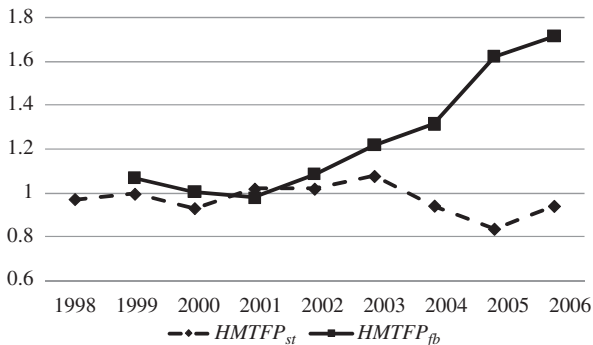


Fig. 5. Static and fixed base HMTFP—case of *BM*.

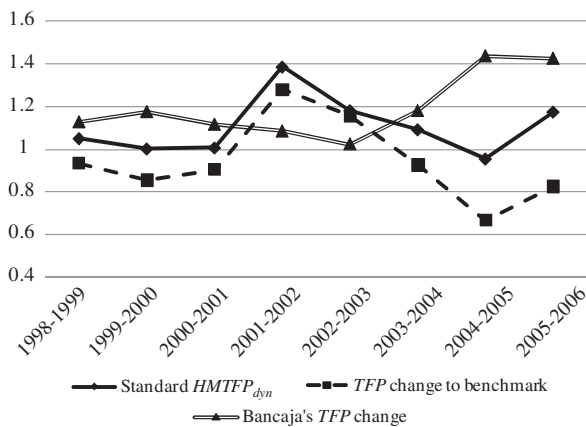


Fig. 6. Dynamic HMTFP and decomposition—case of *BM*.

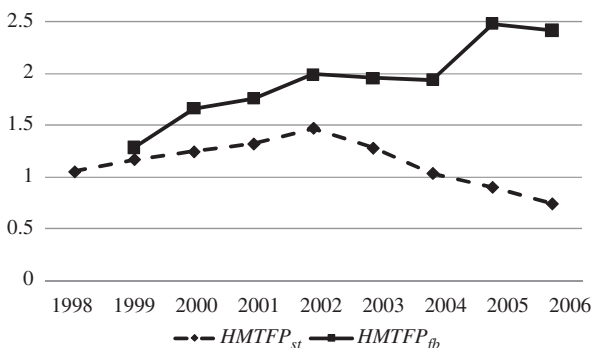


Fig. 7. Static and fixed base HMTFP—case of *GSS*.

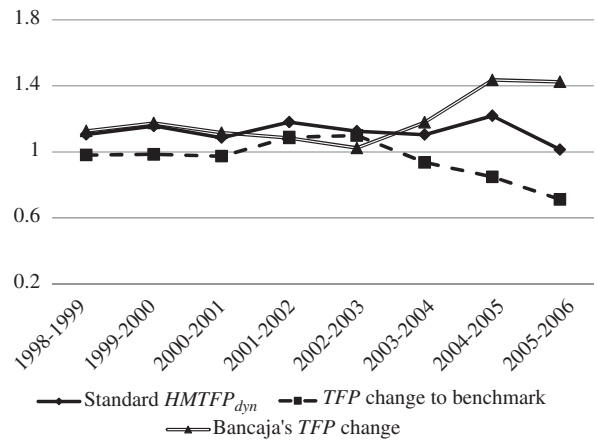


Fig. 8. Dynamic HMTFP decomposition—case of *GSS*.

In Fig. 5, one observes from a static perspective how *BM* performs better than Bancaja only between 2001 and 2003. It is then obvious how the *TFP* distance is quite larger in 2005, with Bancaja performing better. The $HMTFP_{fb}$ index indicates that *BM* also has its highest *TFP* growth in 2005. Until 2002, *BM* has a *TFP* performance close to the fixed base, while higher increases appear starting with 2003. These results should be evaluated jointly with Fig. 6 where the $HMTFP_{dyn}$ relative to the benchmark shows that *BM* has *TFP* changes superior to Bancaja in the same interval as for the $HMTFP_{st}$ index: years 2002 and 2003. However, in the following periods, and similarly to the sample level analysis, Bancaja performs better than *BM*. Finally the standard $HMTFP_{dyn}$ component illustrates that *BM* had its highest *TFP* growth in 2002.

For *GSS*, Fig. 7 shows that from a static perspective this bank is performing better than Bancaja during more time periods, namely 1998–2004. *GSS*'s $HMTFP_{fb}$ index confirms that *TFP* increases are quite high during the first analyzed periods, followed by no growth between 2002 and 2004, and a spectacular increase in 2005. The $HMTFP_{dyn}$ relative to the benchmark (Fig. 8) does not find this same growth, as *GSS* is only superior to Bancaja in 2002 and 2003. Note that these are the same periods as in the case of *BM* (but *GSS*'s scores are closer to the benchmark) and of the sample results, and correspond to the post-merger and expansion years of Bancaja. Concluding, the standard $HMTFP_{dyn}$ index shows that even if in 2005 and 2006 *GSS* has lower *TFP* growth than Bancaja, this growth is still this unit's highest *TFP* improvement.

5. Concluding remarks

This research is founded in the traditional view of benchmarking as the search and emulation of best practices. By applying the *HMTFP* index [19], this study aims at closing the gap between benchmarking and multi inputs and outputs *TFP* frontier analysis. In this way, *TFP* benchmarking can be a new way to set strategic objectives for managers and to analyze firm performance for regulators and researchers.

The advantages of the proposed tool for benchmarking are various. First, this Hicks–Moorsteen type index, which is currently rather scarcely used, solves known problems of *TFP* measurement in the presence of variable returns to scale. Indeed, under weak assumptions of strong disposability and *VRS*, this index is always feasible. This property is crucial for benchmarking analysis as firm-specific results have to be provided. Thus, one implication is that the *HMTFP* index deserves greater attention.

Second, through straightforward manipulations of the *HMTFP* index, versatile tools for benchmarking analysis are obtained.

Pursuing a global image of *TFP* benchmarking, three measures result from diverse assumptions: (1) static benchmark analysis, (2) fixed base and unit benchmark analysis, and (3) dynamic benchmarking analysis using a new decomposition. These benchmarking viewpoints assume fixing a particular *DMU* as a benchmark (very little used in previous analyses) and/or base technologies (a classical benchmark approach) together with the pros of the standard *HMTFP* index. Each of these settings enables managers to see a certain facet of the firm's activity. These benchmarking indices are stand alone tools, but can also be potentially combined to offer a broad perspective for decision making.

For the empirical analysis, this paper used benchmarking criteria based on current interests of banking institutions: technical efficiency and good stress tests results. While technical efficiency has always been considered an important aspect, stress tests attracted a lot of interest in the aftermath of the recent financial crisis. Results confirm the growth phase of the Spanish banking industry and illustrate how *TFP* scores evolve in the sector. This consistent growth phase originated at the end of the deregulation of the savings banks sector. For instance, for these banks Grifell-Tatjé and Lovell [30] found that productivity declined at the end of the 1980s, beginning of the 1990s. For the following period, recent studies find significant productivity increases due to either deregulations [33], or technological change (e.g., [45] or [30]).

Furthermore, the fluctuations encountered in banks' *TFP* may be due to expansion or consolidation strategies, such as mergers. For example, Cuesta and Orea [41] find immediate efficiency decreases for savings banks involved in mergers, followed by significant increases. In our empirical application these fluctuations are revealed through significant catching-up effects. Moreover, this sudden decrease which follows expansion strategies can explain the *TFP* evolution of Bancaja, the benchmark. Also, Bancaja's significant *TFP* growth at the end of the analyzed period is in line with the results of Illueca et al. [36], who state that savings banks that expand outside their original markets achieve greater productivity gains.

Throughout the paper, findings are first scrutinized by comparing the sample level results with the established benchmark under the various benchmarking scenarios. While key results are stable between the different approaches, differences may appear mostly due to the chosen treatment of the fixed or changing technology. Next, the same scenarios are analyzed in a unit to unit analysis revealing dissimilar behaviors of banking units.

This study makes headway for future research since the proposed methodological tools can employ benchmarking criteria adapted to any scenario or industry. For instance, in the standard benchmarking approach comparisons against efficient units reveal the firm's position in the market and the distance separating it from the efficient units. This is a method to discover, understand and implement new organizational practices. In this line it could be interesting to define analyses by benchmarking against strategic groups' leaders. However, in some cases managers may want to compare performance against their local competitor, even if this may be an inefficient firm. For savings banks, this local competitor may be a unit that is developing its branch network in the same region. All these benchmarking options contribute to organizational learning and strategic planning and reveal how decision making can contribute to the performance of firms over time.

Finally a limitation of this study – an avenue of future research – is the absence of risk variables in *TFP* indices. The importance of including risk measures has become acute following the recent financial crisis. One option could be to obtain risk-adjusted estimations of *TFP* (e.g., the work done by Hughes and Mester [50]

in a cost function approach could be adapted to *TFP* indices). For instance, future studies could introduce the risk variables through outputs such as the credit-risk expressed as the amount of off-balance-sheet items. Alternatively one could use banking ratios defining the risk environment (e.g., percentage of insolvency provisions or simply the risk of assets).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.omega.2011.01.001.

References

- [1] Camp RC. Best practice benchmarking: the path to excellence. *CMA Magazine* 1998;72(8):10–5.
- [2] Balk MB. The residual: on monitoring and benchmarking firms, industries, and economies with respect to productivity. *Journal of Productivity Analysis* 2003;20(1):5–47.
- [3] Smith M. Performance measurement and management. London: Sage; 2005.
- [4] Collis D, Young D, Goold M. The size structure and performance of corporate headquarters. *Strategic Management Journal* 2007;28(4):383–405.
- [5] Banker RD, Chang HH, Majumdar SK. A framework for analyzing changes in strategic performance. *Strategic Management Journal* 1996;17(9):693–712.
- [6] Bragg SM. Business ratios and formulas: a comprehensive guide. New Jersey: Wiley; 2002.
- [7] American Productivity Center. Total performance measurement. Texas: Houston; 1981.
- [8] Färe R, Grosskopf S, Lovell CAK. Production frontiers. New York: Cambridge University Press; 1994.
- [9] Ray SC. Data envelopment analysis: theory and techniques for economics and operations research. New York: Cambridge University Press; 2004.
- [10] Färe R, Grosskopf S, Roos P. Malmquist productivity indices: a survey of theory and practice. In: Färe R, Grosskopf S, Russell R, editors. Index numbers: essays in honour of Sten Malmquist. Boston: Kluwer; 1998.
- [11] Fethi MD, Pasiouras F. Assessing bank efficiency and performance with operational research and artificial intelligence techniques: a survey. *European Journal of Operational Research* 2010;204(2):189–98.
- [12] Wheelock DC, Wilson PW. Technical progress, inefficiency, and productivity change in US banking, 1984–1993. *Journal of Money, Credit and Banking* 1999;31(2):212–34.
- [13] Grifell-Tatjé E, Lovell CAK. A note on the Malmquist productivity index. *Economics Letters* 1995;47(2):169–75.
- [14] Briec W, Kerstens K. Infeasibility and directional distance functions with application to the determinateness of the Luenberger productivity indicator. *Journal of Optimization Theory and Applications* 2009;141(1):55–73.
- [15] Glass JC, McKillop DG. A post deregulation analysis of the sources of productivity growth in UK building societies. *Manchester School* 2000;68(3):360–85.
- [16] Yörük BK, Zaim O. Productivity growth in OECD countries: a comparison with Malmquist indices. *Journal of Comparative Economics* 2005;33(2):401–20.
- [17] Cummins JD, Rubio-Misas M. Deregulation, consolidation, and efficiency: evidence from the Spanish insurance industry. *Journal of Money Credit and Banking* 2006;38(2):323–56.
- [18] Kao C. Malmquist productivity index based on common-weights DEA: the case of Taiwan forests after reorganization. *Omega* 2010;38(6):484–91.
- [19] Bjurek H. The Malmquist total factor productivity index. *Scandinavian Journal of Economics* 1996;98(2):303–13.
- [20] Lovell CAK. The decomposition of Malmquist productivity indices. *Journal of Productivity Analysis* 2003;20(3):437–58.
- [21] Briec W, Kerstens K. The Hicks-Moorsteen productivity index satisfies the determinateness axiom. *Manchester School*, forthcoming.
- [22] Bjurek H, Førsund FR, Hjalmarsson L. Malmquist productivity indexes: an empirical comparison. In: Färe R, Grosskopf S, Russell RR, editors. Index numbers: essays in honour of Sten Malmquist. Boston: Kluwer; 1998. [Chapter 5].
- [23] Nemoto J, Goto M. Productivity, efficiency, scale economies, and technical change: a new decomposition analysis of TFP applied to the Japanese prefectures. *Journal of the Japanese and International Economies* 2005;19(4):617–34.
- [24] O'Donnell CJ. Measuring and decomposing agricultural productivity and profitability change. *Australian Journal of Agricultural and Resource Economics* 2010;54(4):527–60.
- [25] Berg SA, Førsund FR, Jansen ES. Malmquist indices of productivity growth during the deregulation of Norwegian banking, 1980–89. *Scandinavian Journal of Economics* 1992;94(Suppl.):211–28.
- [26] Berg SA, Førsund FR, Hjalmarsson L, Suominen M. Banking efficiency in the Nordic countries. *Journal of Banking and Finance* 1993;17(2–3):371–88.
- [27] Zaim O, Färe R, Grosskopf S. An economic approach to achievement and improvement indexes. *Social Indicators Research* 2001;56(1):91–118.
- [28] Färe R, Grosskopf S, Hernandez-Sancho F. Environmental performance: an index number approach. *Resource and Energy Economics* 2004;26(4):343–52.

- [29] Zaim O. Measuring environmental performance of state manufacturing through changes in pollution intensities: a DEA framework. *Ecological Economics* 2004;48(1):37–47.
- [30] Grifell-Tatjé E, Lovell CAK. Deregulation and productivity decline: the case of Spanish saving banks. *European Economic Review* 1996;40(6):1281–303.
- [31] Grifell-Tatjé E, Lovell CAK. The sources of productivity change in Spanish banking. *European Journal of Operational Research* 1997;98(2):364–80.
- [32] Lozano-Vivas A. Profit efficiency for Spanish savings banks. *European Journal of Operational Research* 1997;98(2):381–94.
- [33] Kumbhakar SC, Lozano-Vivas A. Deregulation and productivity: the case of Spanish banks. *Journal of Regulatory Economics* 2005;27(3):331–51.
- [34] Más-Ruiz FJ, Nicolau-Gonzalbez JL, Ruiz-Moreno F. Asymmetric rivalry between strategic groups: response, speed and response and ex-ante versus ex-post competitive interaction in the Spanish bank deposit market. *Strategic Management Journal* 2005;26(8):713–45.
- [35] Tortosa-Ausina E, Grifell-Tatjé E, Armero C, Conesa D. Sensitivity analysis of efficiency and malmquist productivity indices: an application to Spanish savings banks. *European Journal of Operational Research* 2008;184(3):1062–84.
- [36] Illueca M, Pastor JM, Tortosa-Ausina E. The Effects of geographic expansion on the productivity of Spanish savings banks. *Journal of Productivity Analysis* 2009;32(2):119–43.
- [37] Bougnol M-L, Dulá JH, Estellita Lins MP, Moreira da Silva MP. Enhancing standard performance practices with DEA. *Omega* 2010;38(1-2):33–45.
- [38] Avkiran NK, Morita H. Benchmarking firm performance from a multiple-stakeholder perspective with an application to Chinese banking. *Omega* 2010;38(6):501–8.
- [39] Paradi JC, Rouatt S, Zhu H. Two-stage evaluation of bank branch efficiency using data envelopment analysis. *Omega* 2011;39(1):99–109.
- [40] Caves DW, Christensen LR, Diewert WE. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 1982;50(6):1393–414.
- [41] Cuesta RA, Orea L. Mergers and technical efficiency in Spanish savings banks: a stochastic distance function approach. *Journal of Banking and Finance* 2002;26(12):2231–47.
- [42] Berger AN, Humphrey DB. The efficiency of financial institutions: international survey and directions for future research. *European Journal of Operational Research* 1997;98(2):175–212.
- [43] Goddard JA, Molyneux P, Wilson JOS. *European banking: efficiency, technology and growth*. New York: Wiley; 2001.
- [44] Andersen P, Petersen NC. A procedure for ranking efficient units in data envelopment analysis. *Management Science* 1993;39(10):1261–64.
- [45] Wilson PW. Detecting outliers in deterministic nonparametric frontier models with multiple outputs. *Journal of Business and Economic Statistics* 1993;11(3):319–23.
- [46] Prior D, Surroca J. Performance measurement and achievable targets for public hospitals. *Journal of Accounting, Auditing and Finance* 25(4): forthcoming.
- [47] Li Q. Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews* 1996;15(3):261–74.
- [48] Kumar S, Russell R. Technological change, technological catch-up, and capital deepening: relative contributions to growth and convergence. *American Economic Review* 2002;92(3):527–48.
- [49] Sikora M. Timing a big bath to an acquisition, mergers and acquisitions. *Dealmaker's Journal* 1999;33(6):8–9.
- [50] Hughes JP, Mester LJ. Bank capitalization and cost: evidence of scale economies in risk management and signaling. *Review of Economics and Statistics* 1998;80(2):314–25.