Returns to scale of Business Administration research in Germany

Marcel Clermont · Alexander Dirksen · Harald Dyckhoff

Abstract   In order to efficiently allocate academic resources, an awareness of the properties of the underlying production function's returns to scale is of crucial importance. For instance, the question arises as to what extent an expansion of a university department's academic staff would be advisable in order to utilize increasing marginal gains of research production. On the other hand, it is disputable whether an optimal university department size exists. Empirical studies covering these questions render various answers. In this paper, we analyse which properties of returns to scale the Business Administration research of universities in Germany exhibits. On the basis of research data from 2001 until 2009 provided by the Centre for Higher Education, and using Data Envelopment Analysis, we demonstrate that typically sized business schools show nearly constant returns to scale. Furthermore, we observe tendencies of decreasing returns to scale for large-sized business schools. Diverse robustness and sensitivity analyses confirm the validity of our inferred empirical findings.

Keywords   Business Administration research · Centre for Higher Education · Data Envelopment Analysis · Empirical production function · Returns to scale · Scale efficiency

JEL Classification   C61 · D24 · L25

Introduction

An answer to the question of which properties of returns to scale the research and/or teaching technology of universities or university departments feature, is essential for

M. Clermont (✉) · A. Dirksen · H. Dyckhoff
Chair of Business Theory: Sustainable Production and Industrial Control, School of Business and Economics, RWTH Aachen University, Templergraben 64, 52056 Aachen, Germany
e-mail: clermont@lut.rwth-aachen.de

A. Dirksen
e-mail: dirksen@lut.rwth-aachen.de

H. Dyckhoff
e-mail: dyckhoff@lut.rwth-aachen.de

determining an optimal institution size. Such an optimum exists, for example, when returns to scale increase with a growing decision making unit (DMU) size at first and then decrease later. On the minimum efficient scale, the average cost per unit of scientific benefit is minimal (Gutenberg 1983). However, an unambiguous optimum does not always exist, because for example, all DMU sizes are optimal in the event of constant returns to scale. In such cases, the average cost of scientific benefit with constant factor-input ratios and prices is independent of the size of a DMU. That means, for example, that the choice of the particular departmental size is irrelevant to research efficiency in this case, which is different from the case of an unambiguous minimum efficient scale.

Basically, there are plausible arguments for any type of returns to scale in universities or university departments (e.g. Cohen 1980; Kyvik 1995; Fandel 2007). Reasons for increasing returns to scale are, for example, synergy effects induced by economies of scale and scope. Such effects can be explained by potential cooperation and joint research projects as well as more evenly allocated teaching assignments and administrative tasks. Furthermore, there might be a minimum level of successful university research that consists of an adequate number of university departments or corresponding professorships accompanied by a particular infrastructure of personnel and equipment as well as a certain number of students. In contrast, decreasing returns to scale are also conceivable and, for instance, justifiable due to innovation-adverse regulations and routines as well as increased coordination requirements.

Findings of empirical studies that examine correlations between the size of universities or departments on the one hand and research productivity on the other hand do not render unambiguous results with regard to different countries and disciplines (e.g. Johnston 1994; von Tunzelmann et al. 2003). Rather, the authors observe increasing, decreasing, and constant returns to scale. In this respect, Bonaccorsi et al. (2006, p. 393) state that "there is lack of consensus on the existence of economies of scale in scientific production and higher education". Because of the heterogeneity of individual scientific fields, investigations of returns to scale are supposed to be carried out merely for specific scientific disciplines (Abramo et al. 2012). Due to varying organizational structures and research objectives of universities in different countries, such analyses can still only be properly interpreted for institutions within individual countries.

From the point of view of a German Business Administration department or rather of a German business school (hereafter: BuS), Dyckhoff et al. (2009) examine the relation between the size of a BuS and its research output. They show that a BuS comprised of between 8 and 16 professors exhibits approximately constant returns to scale with regard to its research production. Due to the limited number of BuSs with more than 16 professorships, however, the authors are not able to derive results for that respective cluster. They base their study on research data originating from the Centre for Higher Education (CHE) acquired in 2004; i.e. the respective CHE output and input indicators refer to the years 2001 up until 2003.

The non-profit organization CHE conducts comprehensive evaluations of academic research in Germany that differentiates between specific academic disciplines. It acquires and evaluates comprehensive data on German universities' research and teaching activities. Almost every academic discipline is analysed in a 3-year-cycle, in order to create transparency for interested stakeholders. Evaluation results are published on the CHE's website as well as through its media partner, the German magazine DIE ZEIT. CHE data acquisition and evaluation of business research attract great attention in the German business academic community. Especially German governmental bodies and university managements utilize the CHE data and results in order to evaluate research between different BuSs in a

comparative manner. Based upon these assessments, several governmental institutions and universities derive objectives and performance agreements, thus co-opting the results of the CHE research ranking (Clermont and Dirksen 2015).

In the 2004 CHE datasets used by Dyckhoff et al. (2009), research effectiveness of a BuS is represented by three output criteria: the publication points of nationally visible publications, the number of PhD dissertations, and the amount of expended third-party funds. The research efficiency of a BuS is measured by the CHE insofar as each of these output criteria is proportioned to a staff assignment quantity. Since 2004, the CHE data acquisition for BuS research evaluation has been conducted two more times (2007 and 2010), leading to datasets covering the (consecutive) period from 2001 up until 2009. However, the CHE has enhanced the design and data acquisition of several research indicators over the course of time, in order to account more strongly for qualitative aspects of research performance. For instance, the CHE has integrated the number of international publications as an additional output indicator.

Accordingly, the CHE provides an extensive, regularly acquired and enhanced database that is also recognized by the relevant scientific community. As a result, empirical analyses based on these datasets are appropriate. With respect to the analysis of returns to scale, better and more extensive datasets in relation to the research activities of BuSs in Germany are now available through the 2007 and 2010 data acquisitions. This means that analyses regarding properties of returns to scale of German BuSs are now feasible over a consecutive period of 9 years. Thus, from an empirical perspective, the question arises as to whether Dyckhoff et al.'s results of constant returns to scale of German BuS research production represent a robust observation with reference to alternative time periods and evaluation criteria; otherwise it might actually illustrate a one-time result, incurred by the specific design and procedure of data acquisition according to the chosen temporary reference period. Hence, the objective of this paper is to analyse the returns to scale of BuS research production at universities in Germany covering the time period from 2001 up until 2009 and to derive respective implications.

Our paper is structured as follows: In the next section, we provide a brief overview of the state of the art regarding empirical analyses of the properties of research production's returns to scale. Subsequently, we introduce the conceptual design of our examination referring to both the used datasets and the applied methodology. Based upon the described method, we then present and interpret the resulting findings. In addition, we conduct comprehensive robustness and sensitivity analyses to ensure the validity of the inferred results. We conclude our paper by addressing implications and limitations as well as providing future prospects on further research questions.

State of the art

A first overview covering the results of international empirical studies that examine returns to scale in research is provided by the reviews of Cohen (1991), Johnston (1994) and von Tunzelmann et al. (2003). Cohen unconditionally perceives proportional relations between research group size and respective performance, while the other two reviews give more versatile, and even inconsistent, insights. In general, Johnston likewise observes a linear relationship between the size of scientific institutions and corresponding research performance. However, he gives substantial evidence of an optimal institution size existing

(Johnston 1994, p. 32). In contrast, von Tunzelmann et al. (2003, p. 8ff) consider the existence of an optimal institution size in research for groups or teams, but not for (organizationally superior) departments or whole universities.

More current studies disclose heterogeneous results as well, even though most of them constitute a constant ratio between the used input and the produced output (as in Abbot and Doucouliagos 2003; Laband and Lentz 2003; Bonaccorsi and Daraio 2005; Bonaccorsi et al. 2006; Abramo et al. 2012, 2014). Compared to this, Longlong et al. (2009) derive different findings for returns to scale for Chinese universities. They observe either increasing or decreasing returns to scale between a university's size and its research and teaching performances. This, in turn, implies that an optimal institution size of the evaluated universities exists. Van der Wal et al. (2009) and Brandt and Schubert (2013), however, only show decreasing returns to scale with reference to research groups.

With regard to the relation of BuS size in Germany and generated research productivity, so far there mainly exists the study by Dyckhoff et al. (2009). Using Data Envelopment Analysis (DEA), they construct and analyse the empirical production function—generated by the CHE research data of German BuSs between 2001 and 2003—especially with regard to the properties of returns to scale. In view of multi-dimensional total productivity measurements, their examinations demonstrate that BuSs employing between 8 and 16 professors produce research outputs with nearly constant returns to scale. A similar result had already been detected by Ahn et al. (2007), as a minor conclusion while conducting efficiency analyses of the prior CHE research data from 1998 up until 2000.

Altogether, the studies mentioned above mainly differ with respect to the way in which the evaluated organizational units are defined (individual scientists, research groups, departments or universities) and in the choice of indicators. Since we focus our examinations on the relation between the research productivity and the size of German BuSs, we evaluate and extend the results achieved by Dyckhoff et al. (2009). Due to changes in the German university system, it is questionable as to whether their results are valid for subsequent years, too. Since the CHE data acquisition in 2004, the Business Administration research landscape in Germany has changed decisively with regard to numerous influencing factors. First of all, owing to diverse funding programs originating from the federal government and state administrations—for instance, the so called "Excellence Initiative"—many BuSs have grown in size (measured by the number of employed scientists). Moreover, for young scientists, the raising of third-party funds and in particular the placing of articles in international top journals have become an increasingly necessary condition in order for them to be merely considered when they apply for vacant professorships (Schrader and Hennig-Thurau 2009, p. 202). Because of performance-based salary reforms in Germany and performance-based management by objectives, even for already appointed scientists, the pressure to raise funds and to publish research results is increasing (Bort and Schiller-Merkens 2010, p. 340). As the CHE has incorporated the developments outlined above regarding the design and data acquisition of research indicators, we have taken CHE data from 2004 up until 2009 as the basis for our study. In addition, we also use the (already analysed) data from 2001 up until 2003, in order to achieve robust findings with respect to a more comprehensive and consecutive time period.

From a methodological point of view, the previously outlined studies of the state of the art discuss and apply different approaches to examine the properties of returns to scale. What most studies have in common is that they utilize stochastic methods in order to deduce functional relationships between in- and output values of the underlying data (points)—usually by executing one-dimensional and/or multi-dimensional regression analyses. These approaches construe a function that is supposed to reflect the average

transformation of inputs into outputs. Such methods, though, assume no inefficient universities to be existent within the dataset. However, it does not seem plausible that each (considered) university is producing in an efficient manner.

Assuming the generated data to be deterministic as well as comparable to each other, it is possible to analyse dominance relations between the inputs and outputs of the considered universities. This implies inefficient universities to be potentially existent in such a dataset. By definition, a production function only describes the efficient frontier of the underlying production technology (Farrell 1957). Because we focus on properties of returns to scale (with respect to the research production of German BuSs), the efficient frontier (of production)—or more precisely the production function—is our main subject matter. The DEA methodology identifies such a (best-practice) production function by generating an efficient frontier using the underlying input and output data.

In this regard, there are numerous articles in the scientific literature on efficiency analyses of universities that use DEA. A database query on the Web of Science produces 1674 articles in scientific journals based only on the search terms "Data Envelopment Analysis" or "DEA" in connection with "research efficiency" or "teaching efficiency" (as of November 25, 2014). In principle, such articles can be classified based on whether aspects or models of DEA are developed further, whereupon they are validated on the basis of application examples (e.g., Cook and Zhu 2007), or whether the empirical research question regarding the efficiency of universities is emphasized, which is determined on the basis of previously established DEA models (e.g., Fandel 2007). In regard to the empirical applications, the fundamental DEA models are frequently used (e.g., Ahn et al. 1988); however new models, such as DEA applications in hierarchical output structures, or Balanced DEA, are also used for analysing university performance (see Meng et al. 2008; Dyckhoff et al. 2013). Apart from model selection, the applications are also distinguished according to whether teaching (e.g., Johnes 2006; Ray and Jeon 2008), research (e.g., Johnes and Johnes 1993), or both components [integratively (e.g., Tomkins and Green 1988) as well as separately (e.g., Beasley 1995)] are taken into consideration. These applications are further distinguished with respect to the demarcation of the production system, which is to say whether the focal point is that of entire universities (e.g., Breu and Raab 1994) or departments (e.g., Madden et al. 1997; Kao and Hung 2008) or individual professorships or researchers (e.g., Gutierrez 2007). The input and output factors that are used in these articles also vary. In the context of research efficiency analysis, frequently publication and third-party factors are employed (e.g., Johnes and Johnes 1993; Abramo et al. 2011). What all these approaches have in common is that they determine and analyse efficiency scores of the university organizational units that they are investigating.

Analyses of properties of returns to scale with respect to research productivity using DEA, by contrast, is a divergent area of research that, to our knowledge, has only been addressed in a few articles. Given the fact that DEA has been established in efficiency analyses of universities, and Dyckhoff et al. (2009) demonstrated the expedience of DEA with respect to the analyses of properties of returns to scale of German BuSs, we have chosen to utilize this method in our paper. As the description of the methodological approach in the next section will make clear, although our research design differs from that of pure efficiency analyses, the analysis of properties of returns to scale may be helpful in conceptionally designing efficiency analyses using DEA. This is because in most DEA applications, the user determines ex-ante which properties of returns to scale the production system under consideration is featuring. An analysis of returns to scale done in advance could therefore deliver essential informational support for selecting relevant DEA models in following efficiency analyses.

Research design

Datasets: CHE research data on German BuSs

Before analysing the relation between the size of a BuS and its research productivity, first of all we imperatively have to define how the research production of a BuS is actually to be construed and operationalized. In general, quantitatively measurable indicators are used to describe research activities, especially bibliometric measures, such as the number of publications and citations.[1] As already introduced before, a comprehensive evaluation of BuS research production in Germany, and especially the associated acquisition of an approved and accepted database of the research production of German BuSs, is carried out by the non-profit organization CHE. To date, the CHE research data have been available for acquisition years 2001, 2004, 2007 and 2010. The respective data referring to a year of acquisition always correspond to the used inputs and achieved outputs of the preceding 3 years. This means that, e.g., the 2004 dataset represents research production for the years 2001 up until 2003.

Since the data acquisition in 2007, the CHE has operationalized research outputs by using the following four absolute indicators[2]:

- Publication points of nationally visible publications (PP-nat),
- Number of internationally visible publications (Pub-int),
- Number of PhD dissertations (PhD),
- Amount of expended third-party funds (TPF).

The value of the indicator PP-nat is determined via publications of post-doctoral scientists (including professors) of each BuS. For each acquisition period, the BuSs forward lists of post-doctoral scientists' names to the CHE. Thereupon, the CHE identifies the respective publications by querying the literature database WISO, which covers English as well as German literature of Economics and Business Administration. The identified publications are then converted into publication points by incorporating the number of authors, quantity of pages, and, in the case of journal articles, the perceived quality of the publication media—in reference to the journal-ranking JOURQUAL2 of the German Academic Association for Business Research (Schrader and Hennig-Thurau 2009).[3] Since the relevance of journal articles is steadily increasing in the Business Administration discipline in Germany, in 2007 the CHE integrated a new indicator, namely internationally visible publications. The value of this indicator is identified by means of a query on the literature database Web of Science (WoS), which covers scientific literature with a focus on English-language papers. The resulting number of publications, which is acquired by

---

[1] Nevertheless, it has to be considered that depending on the examined academic discipline, varying partial activities, i.e. varying partial indicators, are subsumed under the respective total research production. Research production, in turn, certainly only constitutes one partial aspect of a department's (total) production in favour of a university; especially academic teaching activities remain unconsidered.

[2] On the basis of these absolute indicators, the CHE determines corresponding relative indicators. However, since the exact CHE approach is irrelevant for the subject matter of the present paper, we refrain from an explicit illustration and refer to the appropriate explanations of the CHE (e.g. Berghoff et al. 2011).

[3] However, the quality-weighting for journal articles—in line with the calculation of publication points— had not yet been incorporated into the data acquisition in 2004. Moreover, in 2004 only those publications were included which had at least one professor as (co)author—identified by a BuS-submitted list of all respective professors' names. From 2007 on, however, all publications featuring at least one post-doctoral scientist as (co)author, have been taken into account.

the bibliometric team at Forschungszentrum Jülich, is not converted into weighted publication points, though.[4] The amount of expended third-party funds during an evaluation period and the number of PhD dissertations per semester—covering the six semesters within the 3-year survey period—are acquired by surveying the BuS.

The validity of these preselected CHE output indicators depends on the primary objective of scientific research. This consists in the production of new, usually publicly available, knowledge about the world (Chalmers 1990, p. 23). The dissemination and discussion of this knowledge mainly takes place via written articles. Therefore, indicators based on these publications are generally accepted and well-established in research and practice, in order to evaluate the research of BuSs. By establishing a quality-weighting for journal articles depending on the reputation of the publication medium and by introducing an additional indicator Pub-int (from 2007 onwards), the CHE has furthermore taken additional qualitative aspects into account. Accordingly, publications (in terms of PP-nat and Pub-int) represent appropriate output measures of BuS research.

PhDs per se comply with the objective of educating and training young scientists. Hence, corresponding indicators are initially applicable in order to measure a desired secondary objective of scientific research. Since any PhD degree is inevitably interconnected with the generation and publication of new scientific findings through a published PhD dissertation, the number of PhD dissertations can indeed indicate certain achievements of objectives within the research activities of a BuS.

Compared to the already outlined indicators, the classification of expended third-party funds as a benefit category is less clear-cut, though. Whereas governmental research policy increasingly employs third-party funds as a performance indicator, the scientific literature controversially discusses their applicability for measuring research. This heterogeneous view results from the fact that third-party funds either constitute a resource effort and therefore an input, or they may be interpreted as a desired research achievement—representing a proxy indicator associated with the raising of funds or predictions of their use— and thus an output.[5] It is, then, subject to the respective decision-making context or the underlying objectives as to whether third-party funds are to be regarded as inputs or outputs in research productivity analyses.[6] However, due to their positive perception by governmental research policy in general, as well as by university administrations in particular, the consideration of TPF as a desired output indicator appears to be an acceptable research indicator.

Depending on the examined academic discipline, varying resources constitute the decisive inputs of research activities. But while for several disciplines, production factors such as machinery, laboratories etc. build crucial research inputs, Business Administration research only requires accordant facilities to some extent. Therefore, such factors probably have only little impact on the specific progression of the BuS research production's

---

[4] Relating to Pub-int, there is a difference from the dataset of 2004, too, since this indicator was not part of the CHE indicator set and was not acquired accordingly. Furthermore, due to the lack of Pub-int weighting, there is a certain inhomogeneity between the two publication indicators in both the indicator design and the indicator acquisition. However, because we would like to base our analyses on the original CHE data, we initially forego a conceivable Pub-int weighting analogously to the PP-nat procedure. We address the extent to which such weighting has an impact on the resulting returns to scale in the sensitivity analyses.

[5] A detailed examination concerning the relevance of TPF in research productivity and performance analyses is provided by Rassenhövel (2010, p. 85ff).

[6] Beasley (1990, 1995) and Fandel (2007), for instance, use third-party funds both as an input and an output in their empirical studies.

efficient frontier.[7] In contrast, measures of staff assignment in terms of professors and research assistants and their intrinsic human resources represent the fundamental research inputs. As input factors (and in order to define the size of a BuS), we therefore use the number of (staffed) positions of BuS professors and research assistants. This information is obtained by the CHE directly from the BuSs. According to the CHE surveys, the BuSs were explicitly requested to enter such positions that were financed with state money and not with third party funding.

Consequently, it appears that the CHE indicator design appears quite reasonable from the perspective of the field of German Business Administration. That is also substantiated in the formation and consultation of a scientific advisory board made up of well-known expert German scientists in that particular field. The relevance and acceptance of the CHE data is also apparent in the appreciation and application of this data by German governmental bodies and university managements. The CHE methodology has also been given a positive assessment in the scientific literature (e.g. Tavenas 2004; Usher and Savino 2006; Marginson and van der Welde 2007). For instance, in an analysis by Stolz et al. (2010), examining 25 European data acquisition and subsequent appraisal methods for university evaluation, the CHE is rated best.

To summarize, the CHE has generated core competencies in the acquisition and evaluation of university research indicators since the early 2000s. With respect to the indicator Pub-int, with the bibliometric team at the Forschungszentrum Jülich, the CHE also has expert cooperation partners for matters pertaining to publication indicator design and acquisition from a scientometric perspective. Not least of all, the CHE data acquisitions are carried out continuously and will be continued into the future; this secures the maintenance of current datasets in order to ensure further empirical analyses in the future as well. Even though, there is criticism regarding the CHE indicator design or acquisition in the literature (in general, see, e.g., Frey 2007; Jarwal et al. 2009; Kieser 2012; for a specific focus on the CHE research ranking, see Clermont and Dirksen 2015), we initially accept these disadvantages and analyse which properties of returns to scale German Business Administration research production exhibits when using these preselected research indicators as inputs and outputs. The effects of expedient and feasible modifications to the included research indicators on the validity of our results are investigated within the framework of robustness and sensitivity analyses.

However, the dataset of the first data acquisition in 2001 (inputs and outputs of 1998 up until 2000) can only be considered as a kind of pioneer dataset, used as a foundation for successive enhancements of indicator design and acquisition method in the following years. Due to the acquisition procedure, as well as for validity and consistency reasons, this dataset is applicable for empirical analyses just limited (Gilles 2005, p. 130ff; Dyckhoff et al. 2009, p. 27f). Therefore, considering the years of data acquisition from 2004, 2007 and 2010, research data from 2001 up until 2009 are taken as a basis.

These datasets were provided to us by the CHE and—in terms of internationally visible publications—by the bibliometric team of Forschungszentrum Jülich. The resulting datasets are given in the Appendix, namely Tables 3, 4 and 5. Thereby, only those BuSs are presented that feature complete input and output data. All values embody the total sum of respective absolute indicator values referring to the incorporated 3 years of the survey

---

[7] At the very least, this conclusion is valid for the considered periods of our analyses. Meanwhile however, methodological approaches in Business Administration studies are increasing as well, requiring considerably more physical resources. For example, this applies to neuropsychological and experimental research approaches.

periods. That is, the annualized input and output values are ascertainable by dividing the total values by three. In addition, the top 10 most research-reputable BuSs for each acquisition year are highlighted in bold. The classification of the research reputation of a BuS has been determined on the basis of a supplemental CHE survey. Accordingly, the CHE requested all professors of Business Administration in Germany in each data acquisition year to state up to 5 BuSs that they perceive as being notably strong in research.

Methodological approach: DEA and scale efficiencies

From the perspective of DEA, the BuSs represent DMUs, and the previously discussed input and output indicators constitute their research technology. In analogy to Dyckhoff et al. (2009), we make the following assumptions:

- The considered BuSs are comparable among each other in terms of the respective timeframes investigated, i.e. an absence of untypical BuSs.
- The data are deterministic and unbiased.
- All (conceivable) convex combinations of data points form new fictive, but realizable, activities of virtual BuSs that also belong to the technology of German Business Administration research.

The third above-mentioned assumption corresponds to a data envelopment of all observed data points. The efficient frontier of this five-dimensional (dataset of 2004) or six-dimensional (datasets of 2007 and 2010) convex polyhedron configures the best practice empirical production frontier. In order to determine properties of returns to scale, DEA is an approved instrument, since it has proven to be exceptionally appropriate with reference to multidimensional analyses of the teaching and research efficiency of scientific institutions (see the state of the art). In general, multidimensional productivities are usually subsumed into total productivities; by application of specific weighting factors for all inputs and outputs, aggregation to one-dimensional inputs and outputs and subsequent following quotient creation. DEA enables us—by using the underlying research indicators—to determine the multidimensional relative (empirical) efficient frontier of the research technology as well as to identify its properties. Apart from the multidimensional perspective of DEA, this method's decisive advantage consists of the fact that ex ante no weighting decision of the diverse (research) objectives has to be made, because they are set endogenously. Furthermore, within the scope of DEA, no parametrical production function type has to be presumed a priori, unlike the case of Stochastic Frontier Analysis (SFA).[8]

DEA spans a multidimensional polyhedron, based on the input and output indicators of all analysed BuSs. The efficient frontier of this polyhedron is determined by the group of efficient BuSs, while all remaining (inefficient) BuSs are located inside the polyhedron. Referring to fundamental DEA models,[9] we can generally distinguish between four model types, resulting from two ex-ante user-based basic assumptions. On the one hand, we can postulate either constant or variable returns to scale (CCR vs BCC models). On the other hand, we have to differentiate between input and output orientation. In the case of input-oriented models, the respective efficiency scores are determined by calculating the maximal possible proportional (or radial) input reduction without decreasing any output; and

---

[8] Compared to deterministic DEA, SFA is more applicable in allowing for randomly scattered data. However, we account for this sensitivity problem of DEA with different robustness and sensitivity analyses.

[9] For the basics of fundamental DEA models, cf. in particular the explanations in the pioneer works of Charnes et al. (1978) and Banker et al. (1984).

vice versa with reference to output-oriented models, by computing the maximal achievable output expansion without increasing any input. The method works via equiproportional projection of each considered BuS onto a frontier's facet of the predefined multidimensional polyhedron, which is spanned by linear combinations (CCR model) or convex combinations (BCC model) of efficient BuSs. The larger the distance of a BuS to the efficient frontier, the lower its respective efficiency score is. In the case of an output (input) orientation, for example, an equiproportional increase of each output by 100 % (decrease of each input by 50 %) without increasing any input (decreasing any output) amounts to an efficiency score of 50 %.[10]

In order to obtain insights into properties of returns to scale with reference to the underlying empirically observed production technology, we adopt the concept of scale efficiency. It is defined as the quotient of the CCR-efficiency and the BCC-efficiency score (Banker et al. 1984, p. 1088f). If in a specific size range, CCR-efficiency scores with constant returns to scale equal the respective BCC-efficiency scores with variable returns to scale, the slopes of the empirically determined CCR and BCC production functions are identical. This, in turn, entails that in this size range (of staff assignment), locally constant returns to scale must exist. Thereby, the convex BCC production technology always constitutes a subset of the linear CCR production technology. As a result, the BCC-efficiency scores are at all times higher or equal to the efficiency scores under CCR models. In addition, CCR-efficient BuSs are always BCC-efficient. Therefore, scale efficiencies mandatorily feature values of between 0 and 100 %.

Returns to scale of Business Administration research over time

Analysis of all BuSs

In the context of the following examinations, we conduct separate scale efficiency analyses for the three CHE datasets of the acquisition years 2004, 2007 and 2010. In Fig. 1, the resulting output-oriented scale efficiencies of all BuSs are visualized in dependency on the annualized number of professorships.[11] The data points of 2004 are illustrated as rhombuses, data points of 2007 as triangles and those of 2010 as circles.[12] The 10 BuSs

---

[10] For further basic information on DEA and corresponding mathematical model descriptions, cf. e.g., Cooper et al. (2007).

[11] First, all deliberations refer to results based on output-oriented DEA models. From a short-term and middle-term point of view, this proves to be plausible, since scientific staff can generally be considered as (almost) fixed in the short run and rather variable in the long run. Accordingly, input-oriented models only make sense for long-term considerations—if at all. Within the scope of the sensitivity analyses later on, we will also address results originating from input-oriented models.

[12] Due to the predefined CHE indicator set, the presented scale efficiencies of 2004 are still based on three outputs (excluding Pub-int). While the calculated efficiency scores in the context of DEA can only increase with the addition of supplementary indicators, scale efficiencies as a quotient of BCC and CCR efficiency score can either increase or decrease. This may basically limit the intertemporal comparability of our scale efficiency analyses. However, our sensitivity analyses indicate that the results and the statements are robust or valid with respect to modifications to the preference relation, which is to say that the exclusion of the Pub-int indicator from the 2007 and 2010 datasets does not produce any notable changes in our results.
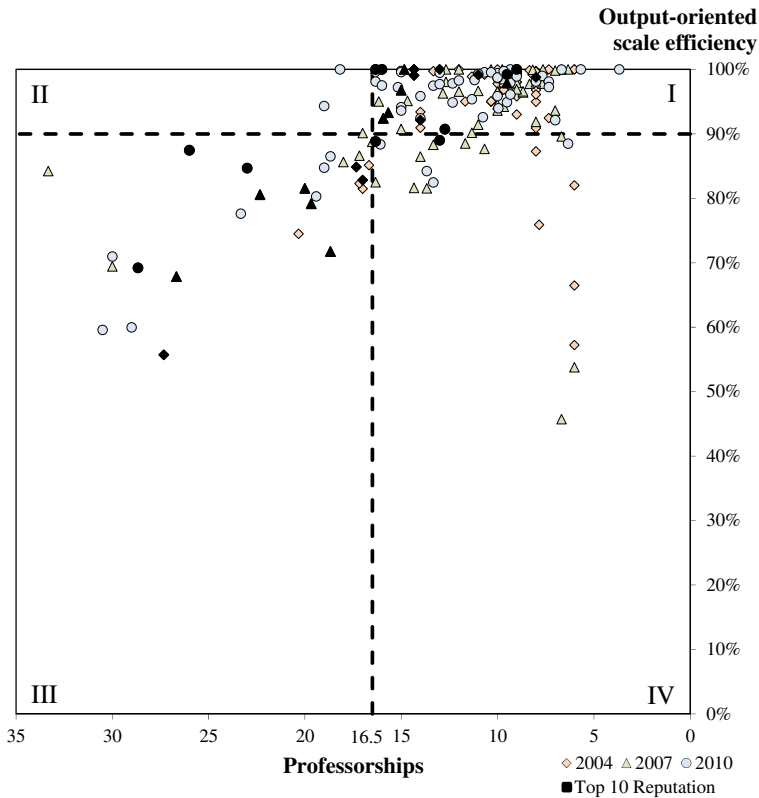
Fig. 1 Output-oriented scale efficiencies according to the number of professorships over the three time periods

featuring the highest research reputation for the respective acquisition year are highlighted in black.

It becomes evident that for all three periods, only a few BuSs employ under 8 or more than 16.5 professors. But whereas the number of small BuSs (under 8 professors) remains constantly low over time, the quantity of larger BuSs (more than 16.5 professors) rises in the course of the periods, and a large number of reputable BuSs are among this group. Based on the dataset of 2004, Dyckhoff et al. (2009) discovered approximately constant returns to scale for a size range of 8–16.5 professors. In this range, all BuSs are characterized by output-oriented scale efficiencies of at least 90 %, most even up to nearly 100 %. To facilitate a comparison to these findings, in Fig. 1 two dotted lines divide the diagram into four sectors (I–IV). Hence, one line indicates a scale efficiency of 90 % and the other marks the (previously) identified number of 16.5 employed professors. When considering all three periods, it becomes apparent that some BuSs in the size range of 8 up to 16.5 professors are located below scale efficiencies of 90 %. Therefore, these BuSs prevent an analogous (unrestricted) transfer of the previously discovered empirical conclusion (of constant returns to scale) onto the data of 2007 and 2010.[13] But remarkably, the

---

[13] If, however, one generously operationalizes locally constant returns to scale already from a scale efficiency of ≥80 %, a size range results from 8 up to 18.6 professors, in which over all timeframes investigated

majority of all BuSs are located within the sectors I and III, while sector II only contains two BuSs (2010) and sector IV merely a few, especially smaller BuSs (of 2007 and 2010). When further taking the respective ten most reputable BuSs of each acquisition year as a basis—thus accounting for the expertise of the evaluated professors with regard to the research-strongest BuSs in Germany—then these are almost entirely located within the sectors I and III.

Generally, most BuSs lie within sector I. All inefficient BuSs in this sector are radially projected onto edges of the BCC-efficient frontier. Thereby, the BuS-specific output mix determines the concrete direction of the projection. The (virtual) outputs of those edges, in turn, must match the outputs of the corresponding reference point on the CCR-efficient frontier by at least 90 %. This also means that the distance between the BCC-production and CCR-production function in this region merely amounts to 10 % of the outputs at maximum (in the case of an equiproportional projection). Accordingly, these BuSs could raise their efficiency score—presuming constant returns to scale—only from 90 to 100 % at most, hence by roughly 11 %. The assumption of (approximately) constant returns to scale in this size range seems plausible at this point.

In contrast, the larger sized BuSs in sector III exhibit smaller scale efficiencies. BCC-frontiers and CCR-frontiers differ quite substantially from one another, which exhibits an indication of decreasing returns to scale for larger sized BuSs. Indeed, such conclusions must generally be stated conditionally, because there are only a few BuSs located in sector III. But certainly a high fraction of these consists of highly research-reputable BuSs; precisely those that are considered to be "strong in research" by the surveyed experts. This is why this observation has potentially higher validity.

In the considerations and results presented above, only the input factor professorships has been used so far to define the size of a BuS. However, research assistants also make a crucial contribution to the research production of a BuS, which is why the CHE acquires them as well and also why they are integrated into our DEA analyses as a second input factor. In order to analyse what influence the number of research assistants has on the respective scale efficiency and on the resulting returns to scale, these two input factors are compared in Fig. 2. The number of professorships is still on the abscissa and the number of research assistants is now on the ordinate. In Fig. 2, we abstract from the precise degree of the scale efficiency by only distinguishing between scale efficiencies that are greater than or equal to 90 % (grey) and scale efficiencies that are less than 90 % (white).

An apparently linear relationship between Business Administration professorships and research assistants becomes clear, which is reflected in relatively high correlation coefficients (between 0.62 and 0.75, depending on the acquisition year, with a significance level of 1 %). Although these two inputs do not have a perfect linear relationship, they rather correspond to each other in a certain range. We can infer from Fig. 2 that the assertion of constant returns to scale with respect to professors is independent of the number of research assistants assigned to them. By contrast, we cannot derive clear assertions about properties of returns to scale only with respect to the number of research assistants.[14] On that basis, the number of research assistants is not included in the definition of the size of a BuS in the following analyses.

Footnote 13 continued

approximately constant returns to scale can be discovered. For the following analyses, though, we will define such a size range of constant returns to scale more closely, in order to assure a higher robustness of our results.

[14] Note that the number of research assistants is, however, always part of our scale efficiency analyses.
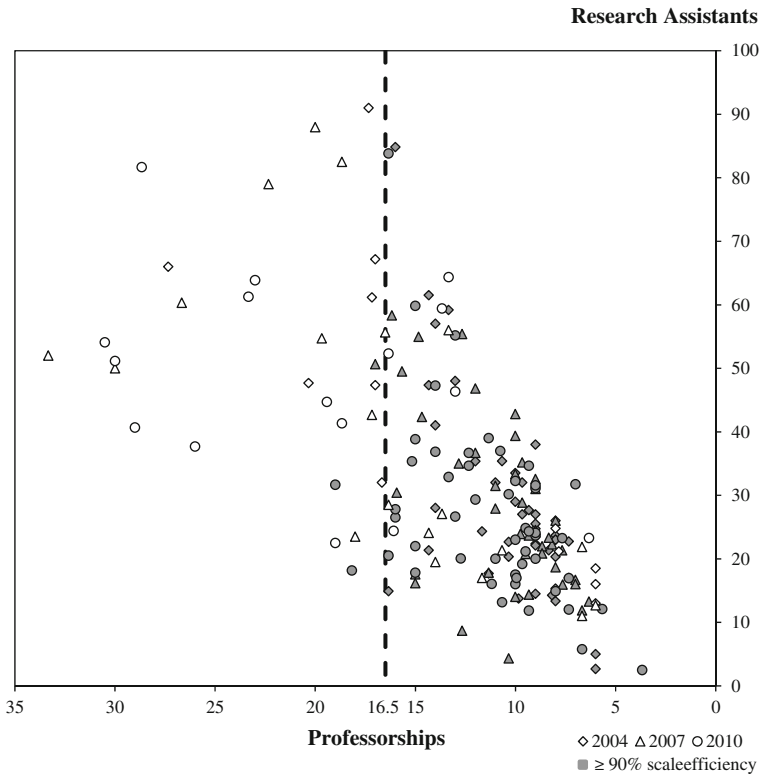
Fig. 2 Relationship between professorships, research assistants and scale efficiency

In order to analyse the aforementioned conclusions of constant returns to scale for medium-sized BuSs and decreasing returns to scale for larger-sized BuSs in more detail, Fig. 3 illustrates the output-oriented scale efficiencies, dependent on the number of professors for each of the acquisition years separately. Due to their limited number, conclusions with respect to smaller BuSs are necessarily problematic. In contrast, deriving findings with reference to the group of larger BuSs is rather reasonable, because of this group's quantitative increase over time as well as the major presence of reputable BuSs. Based on these deliberations, we subsequently take a closer look at the respective size range of between 8 and 28.7[15] professors (reputable BuSs are again highlighted in black). Next, in order to mitigate the impact of single data outliers on an overall conclusion, we approximate the calculated data points with a linear regression assuming a quadratic function.[16] In Fig. 3, the resulting functional relations are illustrated for each of the three timeframes.

If we now operationalize approximately constant returns to scale by means of scale efficiencies of more than 90 %, then the point of intersection between regression and 90 % scale efficiency determines the crucial transition region. It marks the annualized number of

---

[15] This number (of professors) corresponds to the largest of all top 10 reputable BuSs over the three consecutive timeframes investigated.

[16] We applied the respective R^2 as the quality criterion for the function approximations. With values of 0.753 (2004), 0.632 (2007) and 0.545 (2010), they all exhibit high levels.
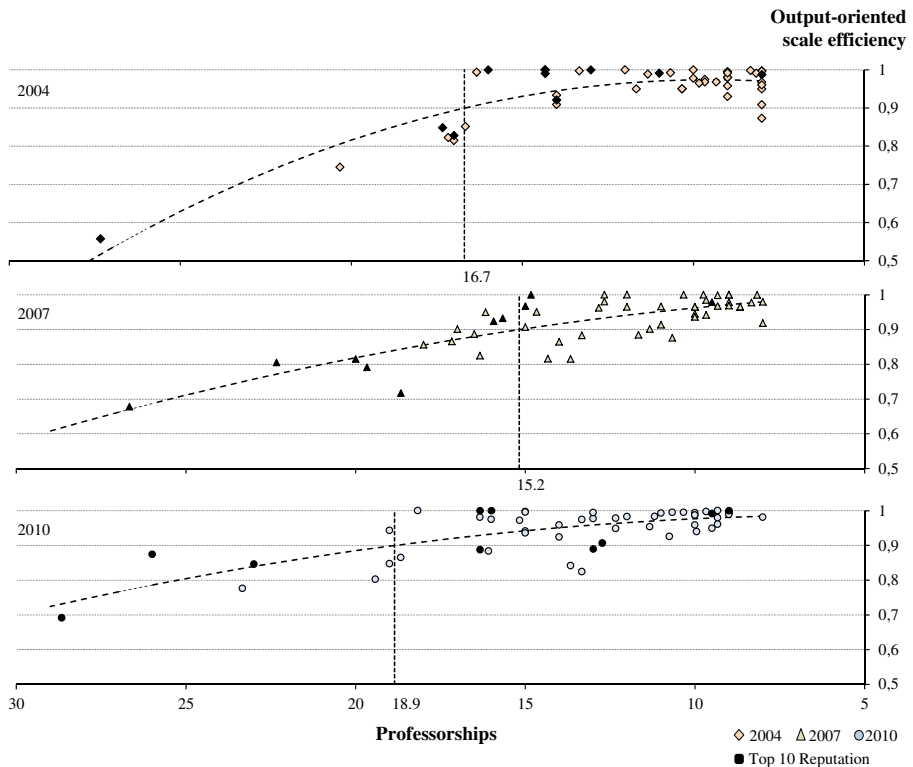
Fig. 3 Regressions of the output-oriented scale efficiencies according to the number of professorships over the three time periods

employed professors up to which the assumption of constant returns to scale can be confirmed in terms of the respective timeframe. Supporting the results of Dyckhoff et al. (2009) for 2004, we identify a size range from 8 up to 16.7 professors of nearly constant returns to scale. This size range slightly decreases for the timeframe of 2007 (from 8 up to 15.2 professors), whereas it increases for 2010 (from 8 up to 18.9 professors).[17]

Because of the aforementioned presumption of convexity with regard to the data envelopment (in the context of BCC-production functions), approximately locally constant returns to scale in the respective size ranges ceteris paribus imply that smaller BuSs consisting of less than 8 professors must exhibit increasing, and larger BuSs with more than 16.7 (2004), 15.2 (2007) or 18.9 (2010) decreasing, returns to scale. With reference to practical applications, such observed size ranges of constant or approximately constant returns to scale are typically rather small; in consideration of many medium-sized, but only comparably few small and large DMUs. Hence, the relatively large size ranges of approximately constant returns to scale in the course of the consecutive timeframes investigated turn out to be remarkable.

---

[17] When utilizing all data points for each of the timeframes investigated instead of the interval from 8 up to 28.7 professors focused on here, we observe comparable results. However, the approximation of data points by regression falls off in quality, as measured and revealed by lower R^2 values.

As mentioned before, drawing conclusions about the properties of returns to scale with regard to larger BuSs for the acquisition year 2004 is inevitably problematic because of their limited number. In view of 2007 and 2010, however, particularly the number of larger BuSs is increasing substantially, which enables us to derive statements in favour of returns to scale. Due to decreasing marginal scale efficiencies of the regressions in Fig. 2, i.e. lower scale efficiencies of larger BuSs, the tendency of decreasing returns to scale with regard to (too) large BuSs appears to be confirmed. Furthermore, an optimal BuS size with respect to research productivity seems to exist; from this size onwards, marginal gains of additional professorships are diminishing. In order to examine the last-mentioned conclusions more precisely within the scope of an individual analysis of particular BuSs, in the next section we examine the temporal development of research-reputable BuSs' returns to scale. A more detailed analysis of reputable BuSs based on expert judgments by the scientific community for Business Administration in Germany on their own peer-group should reveal interesting and significant results.

Analysis of research-reputable BuSs

Figure 4 illustrates an extract of Fig. 1, focusing on intertemporal movement patterns of scale efficiencies with reference to seven reputable BuSs. In particular, all those BuSs are considered that are part of the top 10 most research-reputable BuSs in each of the three consecutive acquisition years. The depicted paths indicate the temporal development of these BuSs with respect to both their size and (resulting) scale efficiency over the three acquisition years.

The movement patterns confirm the results previously discussed. On the one hand, some reputable BuSs are located within sector I of constant returns to scale over all time periods; undertaken moderate expansions or reductions of employed professors do not implicate their scale efficiencies to be decreasing to such an extent that they would move outside of this sector. On the other hand, for other reputable BuSs, more considerable expansions are accompanied by a reduction of the respective scale efficiencies, especially when exceeding 16.5 professors. For instance, BuS 55 (*Münster*) is steadily increasing in size, which is accompanied by a slight decrease in scale efficiency from 100 (2004) to 93 % (2007) and finally to 89 % (2010). A similar result can be deduced by observing BuS 70 (*Vallendar*). While this BuS is still scale-efficient in 2004 and 2007, after a considerable increase in the number of professors, the scale efficiency is reduced to a mere 87 % in 2010. This trend becomes particularly evident for BuS 50 (*Mannheim*), which is the most research-reputable BuS in Germany over all three surveys. Still scale-efficient in 2004, its continuous enlargement is also accompanied by a decrease in scale efficiency to 81 % (2007) and ultimately 69 % (2010). In contrast, referring to BuS 52 (*LMU München*)—which is the second most reputable BuS in Germany in all three surveys—we can observe an opposite movement pattern. Starting from low scale efficiencies in 2004 and 2007, it achieves a scale efficiency of 100 % in 2010, while simultaneously reducing its professorships.[18]

---

[18] It should be pointed out here that the meaning of a scale efficiency is not comparable with the efficiency score of an original productivity analysis using DEA. A high scale efficiency indicates that the DMU is located in a size range in which the efficient frontiers of the production functions, that are empirically created using DEA by assuming constant and variable returns to scale, are close together. This value does not allow for the derivation of any assertions about the research efficiency of an individual BuS, which is why no concrete efficiency improvement recommendations can be made. Such recommendations rather depend on the specific system of objectives of a BuS and are not part of our paper.
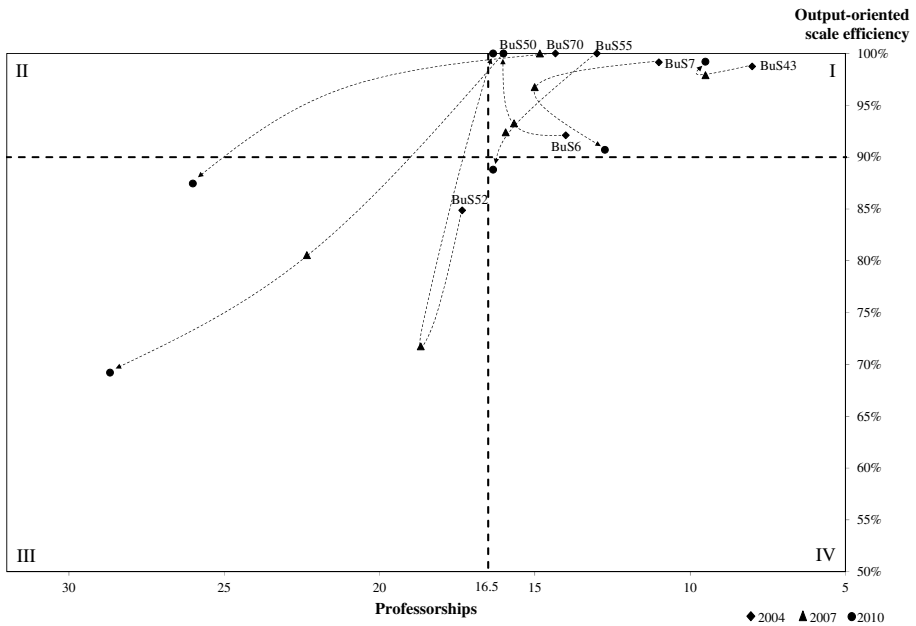
Fig. 4 Movement patterns of reputable BuSs according to output-oriented scale efficiencies and number of professorships over the three time periods

Thus, altogether we can claim that the considered reputable BuSs, by trend, grow in size over time, accompanied by decreasing scale efficiencies. Moreover, it should be noted in particular that in this section we have exclusively investigated movement patterns of research-reputable BuSs. Because of the high research reputation assigned to each of them, we obviously cannot imply that they are particularly unproductive; this, in turn, gives greater weight to our observation. Accordingly, both previously derived main findings—regarding the size ranges of constant returns to scale and decreasing returns to scale afterwards—can still be confirmed subsequent to this detailed investigation of reputable BuS-specific temporal movement patterns.

Robustness and sensitivity analyses

The observed empirical findings are valid in line with the ex-ante chosen framework of production theory, and with the underlying datasets and allowing for the assumptions set. In particular, the present examinations of returns to scale are theoretically founded on production theory and based on the fundamental causality between production function and the concept of efficiency, and subsequently scale efficiency. Efficiency is always relative in a twofold sense: First, it depends on the underlying preference relation—i.e. the choice of considered performance indicators and their attributed values—as well as the existing spanned production possibility set—i.e. potentially realizable activities in the form of real or virtual BuSs. Hence, evaluation modifications in these two criteria have an impact on the derived production function and the corresponding properties of returns to scale. In the

following, we therefore analyse the robustness of the obtained results using sensitivity analyses with respect to these two criteria.

Sensitivity analyses with respect to the production possibility set

When selecting DMUs in the context of productivity and efficiency analyses, their comparability constitutes an elementary prerequisite. Only under the assumption that production technologies as well as considered inputs and outputs of the respected DMUs are homogeneous—i.e. they pursue the same objectives—do productivity analyses within this peer group make sense at all. Since we base our analyses on departments of the same academic discipline, a certain degree of homogeneity appears to be ensured. But because each BuS is integrated into a university-specific environment that is characterized by heterogeneous external parameters and differing stakeholder interests, resulting in individual department structures and possibly diverse objectives, the existence of deviations is quite plausible.[19] In order to reveal such characteristics, we now abstract from BuSs turning out to be untypical and subsequently analyse the resulting modified efficiency scores.

When selectively analysing certain partial productivities—meaning exclusively focusing on each input/output combination—the respective efficient frontiers of the determined partial production functions are usually substantially characterized by just a few BuSs featuring exceptionally high outputs or low inputs. Single erroneous or untypical data points on the efficient frontier exert a potentially strong influence on (in-) efficiency scores of other BuSs and thus also on the results regarding properties of returns to scale. In contrast, inefficient BuSs—having no impact on the efficient frontier—are only responsible for their own efficiency scores. Therefore, we focus on efficient outliers. When closely observing (all) partial productivities for each acquisition year, several BuSs are conspicuous, outperforming the others considerably with reference to their total value. Hence, they are taken out of the datasets in the scope of the following sensitivity analyses.

The process is exemplified in Fig. 5. This figure illustrates the partial productivities of the BuSs with respect to the internationally visible publications (ordinate) and the number of professorships (abscissa) for 2010. It becomes apparent that BuS 52 is clearly distinct from the other BuSs with regard to this partial productivity and especially with respect to the absolute number. Hence, this BuS substantially determines the respective partial efficient BCC-frontier of the empirically identified partial production function and thus may have an influence on multidimensional efficiency analyses using DEA. Hence, BuS 52 is identified as an outlier and is excluded in the following.

Similar arguments apply to BuS 70 in 2004 concerning the indicator PP-nat, to BuSs 53, 56 and 71 in 2007 with reference to TPF as well as to BuS 56 in 2010 regarding PhD dissertations. All of these six BuSs are BCC-efficient; except for BuSs 53 and 56 in 2007, they are also CCR-efficient[20]; beyond that, all six are benchmarking-partner for other BuSs multiple times. So, they might have (considerably) influenced the previously determined

---

[19] Utilizing DEA certainly involves the advantage that respective weightings of objectives are determined (endogenously) for each BuS individually and most favourably. In this way, taking into account BuS-specific profiles is per se inherent in DEA analyses. Thus, the problem of deviating BuS objectives is already mitigated ex-ante by the method itself.

[20] These two BuSs are not CCR-efficient, merely because BuS 71 exhibits the highest partial efficiency with regard to TPF in 2007; BuS 71 therefore obviously has to represent one benchmarking partner for both of them. However, this does not change the fact that all three BuSs feature untypically high TPF values and thus have to be considered in the scope of sensitivity analyses.

Fig. 5 Two-dimensional partial productivity referring to the indicator Pub-int 2010
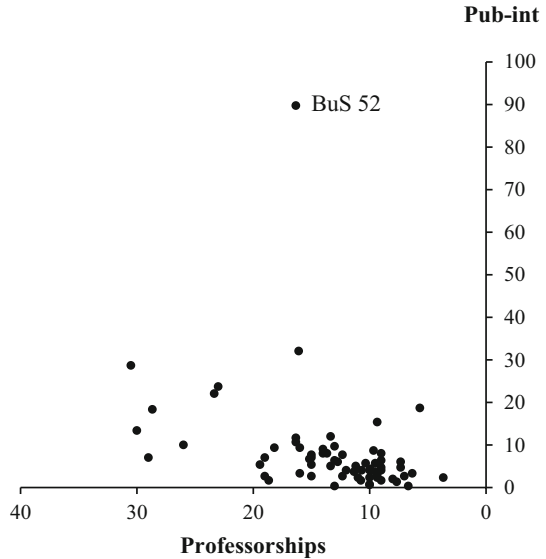
properties of returns to scale. If we now classify these six BuSs as untypical and exclude them from our DEA analyses, in fact the single efficiency scores increase, while the corresponding output-oriented scale efficiencies still remain on a very high level (see Table 1).

When comparing the resulting (modified) scale efficiencies with the original calculations and especially the identified range of constant returns to scale, then the scale efficiencies of 2004 and 2007 barely differ from one another. Only in 2010 do they decline more (by 5 %), but still exhibit 91.2 % on average. Therefore, the conclusion of constant returns to scale for medium-sized BuSs seems to be still valid. The differences in scale efficiency levels of all BuSs in comparison to those from the interval of constant returns to scale primarily result from (the described) strongly decreasing scale efficiencies of larger BuSs; thus, our previously deduced conclusion of decreasing returns to scale for large BuSs can still be verified.

Sensitivity analyses with respect to the preference relation

In order to assess the influence of individual research outputs on the properties of returns to scale, in the following we vary the comprehensive set of partial output components. For this purpose, in each case the indicator Pub-int together with one further output indicator is excluded from the respective DEA analysis.[21] As can be seen in Table 2, excluding certain

---

[21] We intentionally decided to exclude the indicator Pub-int from all sensitivity analyses with respect to the preference relation, since this indicator was not raised in 2004. However, for the datasets of 2007 and 2010, we calculated the accordant scale efficiencies both including and excluding Pub-int, respectively. When comparing the resulting scale efficiencies of the cases a)—excluding just one indicator—and b)—excluding this indicator and additionally Pub-int—, then the resulting regression line adds up to $y = 0.9469x + 0.0405$ ($R^2 = 0.8181$) in the event of additionally excluding PP-nat, $y = 0.9889x + 0.004$ ($R^2 = 0.9166$) in the case of additionally excluding TPF, and $y = 1.0133x - 0.0196$ ($R^2 = 0.9126$) if additionally excluding PhD dissertations. According to this, the obtained results are nearly identical, which is why we only illustrate case a) in our sensitivity analysis of this subsection.

Table 1 Averaged output-oriented scale efficiencies in the case of untypical BuSs being excluded

| | Averaged output-oriented scale efficiency | | | | | |
| | 2004 | | 2007 | | 2010 | |
	All BuSs (%)	8–16.7 professors (%)	All BuSs (%)	8–15.2 professors (%)	All BuSs (%)	8–18.9 professors (%)
Considering all BuSs in DEA	92.9	96.9	91.1	95.1	93.4	96.2
Exclusion of untypical BuSs in DEA	91.9	95.6	91.1	94.7	88.2	91.2

Table 2 Averaged output-oriented scale efficiencies in the case of the indicator mix being varied

| | Averaged output-oriented scale efficiency | | | | | |
| | 2004 | | 2007 | | 2010 | |
	All BuSs (%)	8–16.7 professors (%)	All BuSs (%)	8–15.2 professors (%)	All BuSs (%)	8–18.9 professors (%)
Original indicator set	92.9	96.9	91.1	95.1	93.4	96.2
Exclusion of Pub-int	–	–	90.2	94.0	92.9	95.9
Exclusion of Pub-int and PP-nat	92.8	96.3	87.4	94.7	87.4	91.9
Exclusion of Pub-int and TPF	92.1	96.2	89.6	93.7	93.9	96.5
Exclusion of Pub-int and PhDs	89.9	95.3	89.7	93.8	93.4	96.5
Pub-int weighted	–	–	90.6	94.6	92.1	95.2

output indicators only exerts minor influence on our observed empirical results. In the previously identified size intervals of constant returns to scale, the corresponding averaged four-dimensional output-oriented scale efficiencies are—over all sensitivity analyses and timeframes investigated—considerably higher than 90 %. The low scale efficiency values considering all BuSs are caused by (partly strongly) decreasing scale efficiencies of larger BuSs as well. However, it is quite obvious that the averaged scale efficiencies in the respective size range of constant returns to scale are decreasing distinctly in the course of the investigated periods when both indicators addressing publication activities (Pub-int and PP-nat) are excluded simultaneously. As a consequence, with 91.9 % in 2010, they feature—compared to all other partial considerations—the lowest values in the 'constant' interval. Hence, the results tend to react more sensitively to a simultaneous exclusion of both publication indicators.

As noted in the description of the indicator design and acquisition before, the nationally visible publications are weighted based on the number of authors and the perceived quality of the publication medium. By contrast, such weighting is not done for the internationally visible publications by the CHE. For the benefit of a symmetrical procedure, we therefore also investigate whether differences result concerning the properties of returns to scale when the internationally visible publications are weighted analogously to the nationally visible

publications. Thereby, we weight the number of authors in accordance with the process used in the CHE's PP-nat. With respect to journals' quality weighting, we do not refer to the JOURQUAL2 but rather use the citations of the respective articles. In doing so, we take the divergent acquisition methodology of the Pub-int indicator based on a WoS query into consideration. In general, the impact on scientific advancements and the dissemination of knowledge is operationalized based on citations to create conditions for knowledge spillover benefits. Citations thus represent a proxy measure of the value of output and are therefore predestined for a corresponding weighting. However, the citations of the articles compiled in the Pub-int indicator follow a strongly skewed distribution, and numerous articles remain uncited at the moment of inquiry. For that reason, the actual citations are not used for weighting here. The WoS journal impact factor offers an alternative weighting option, particularly with respect to journal quality. We therefore weight each article in the internationally visible publications with the WoS journal impact factor at the respective moment of inquiry.[22] As Table 2 shows ("Pub-int weighted" row), however, these modifications have no decisive impact on the revealed properties of returns to scale, which is why our prior assertions also apply to a corresponding weighting of internationally visible publications.

In addition, we analyse the effect on the properties of returns to scale, if we regard third-party funds as a research input instead of an output. In this case, the averaged output-oriented scale efficiencies still exhibit high values and show similar characteristics to those already identified in the scope of the previous sensitivity analyses.

The respective progressions of the function approximations—calculated analogously as before—also confirm this observation. Hence, similar intersections are to be seen compared to Fig. 2, thereby the size range considered to be constant ceteris paribus often diminishes slightly. Only an exclusion of both publication indicators leads to considerable shifts of intersections. Accordingly, the previously observed plausible sensitivity and relevance of the two publication indicators are confirmed once again.

One further (possible) modification addresses the orientation of the employed DEA model. Presuming an input orientation in the context of scale efficiency measurements via DEA leads to a less distinct revelation of the empirical findings presented and discussed in the previous sections. Due to the conspicuously weak output performance of many BuSs regarding the CHE postulated inputs, this is no effect, however, that is empirically rich in content. Numerous BuSs feature relatively low output indicator values in this case. Reducing personnel inputs, equiproportionally, while maintaining their low levels of output performance, BuSs are projected onto straight-line segments of the BCC-efficient frontier consisting of less than 8 professors and thus feature, in some cases strongly, increasing returns to scale. This, however, does not enable any implications concerning other segments of the BCC production function consisting of more than 8 professors.

Implications, limitations and future prospects

On the basis of the non-parametric method of DEA, we analyse which properties of returns to scale German BuSs exhibit with reference to research production for the nine-year time

---

[22] The adequacy of using journal impact factors to quantify the outcome of research production is certainly subject to the specific context of application. In this respect, the use of impact factors instead of citations can be accepted for the specific subject matter of Business Administration research, where actually numerous articles are uncited. It is, of course, not valid for a general theoretical perspective in terms of a generally accepted approach for all academic disciplines.

period from 2001 until 2009, thus evaluating and extending the analyses by Dyckhoff et al. (2009) for the period 2001 to 2003. We employ the CHE research data to generate valid datasets of relevant input and output criteria. Despite high dynamics in terms of the number of included BuSs, their sizes as well as the design and acquisition of evaluation criteria, we can determine clear and robust results. We discover that over all considered consecutive periods, medium-sized BuSs from 8 up to 16.7 (2004), up to 15.2 (2007) and up to 18.9 (2010) averaged annualized professorships feature approximately constant returns to scale in research. For larger BuSs above these size ranges, returns to scale decrease (quite strongly in some cases), indicating that BuSs should not increase their academic staff in excess of a certain (critical) level—at least from a research productivity point of view. This, in turn, implies an optimal (productivity-) size of BuSs to exist in a medium-sized range. By means of diverse sensitivity analyses, we illustrate that our findings are also valid for different modifications of influencing parameters.

Certainly, our study is subject to some limitations, which should be reflected on. They are either of an empirical nature, due to utilization of the CHE data, or of a methodical kind, due to application of the DEA methodology. Apart from the discussed general deliberations regarding input and output indicator characteristics and adequacy, the CHE design and acquisition should be scrutinized. For instance, biases of data may occur, because the BuSs classify their Business Administration professorships by themselves. This is due to the fact that the definition and the differentiation of Business Administration professorships from their academic environment are not distinct. The intertemporally varying numbers of professors reported by accordant BuSs, suggest that the differentiation from related and overlapping academic disciplines, e.g. Business Informatics, might not always have been executed in a consistent manner (for a detailed analysis from an accounting perspective, cf. Clermont and Dirksen 2015).

For output indicators, data biases might also occur. This especially applies to the acquisition of publication data using queries in literature databases. In their studies, Clermont and Dyckhoff (2012a, b) reveal that none of the common literature databases feature a complete coverage of all relevant business and economic journals for German Business Administration researchers. But altogether, against the background of practical implementation, the considered indicators seem acceptable under the given conditions; in particular by reasons of diverse and continuously improved procedures of quality assurance by the CHE. In this way, the distinctiveness and robustness of our empirical results becomes particularly evident (especially over the different consecutive timeframes).

Moreover, the research situations of BuSs are characterized by diverging environmental surroundings and workloads in teaching (for instance, public universities versus private universities), which are not considered by the underlying datasets. Therefore, we ultimately conduct partial productivity analyses of the comprehensive academic productivity of BuSs per se, even in the present multidimensional case. Interdependencies between teaching and research activities are neither taken into account. Thereby, basically both positive and negative effects of teaching activities on research productivity are plausible. In this sense, teaching can be considered to be a time-consuming endeavour on the one hand; but on the other hand, discussions with students might also promote research activities.

Further possible limitations result from the applied analysis method of DEA. By reasons of the postulated property of convexity regarding the datasets (within the limits of generating the BCC production functions), the empirical production functions are inevitably concave, so that returns to scale can only be increasing for low input quantities and decreasing for large quantities. Hence, for a significant interval of constant returns to scale to be deducible at all, datasets particularly have to include small and large BuSs amongst

the BCC-efficient ones, adequately positioned in multidimensional data envelopment. For this reason, detected size ranges of constant or nearly constant returns to scale are typically less extensive in practical applications if there are many medium-sized but comparatively few small and large DMUs. Therefore, referring to our empirical findings, the relative extensive size range of (approximately) constant returns to scale over a period of 9 years is exceptionally remarkable. For this interval as well as the observation of decreasing returns to scale for larger BuSs, our results are quite reliable, while statements concerning the properties of returns to scale for smaller BuSs are problematic due to their limited number.

Conditioned by the defined production possibility set and the chosen preference relation, both of which are dependent on the available (and used) data, as well as the diverging interests of different stakeholders (or decision-makers), the informative value of this study is limited with regard to concrete recommendations for policy decisions. An unreflected adoption of our findings in order to justify, for example funding decisions for universities in Germany might then result in incorrect or misguided incentives. Hence, our study represents a partial (specific) perspective of just one area of a university's tasks, that is to say research. Furthermore, not all thinkable and/or possible facets of research production are covered. Therefore, our study should certainly not serve as an exclusive basis for decision-making.

Since there is a lack of specific theoretically founded explanations, we can merely speculate about the reasons for our empirical findings. Accordingly, substantial research synergies in the form of economies of scale and scope between German Business Administration professorships within the same BuS appear to be hardly existent. On the contrary, administrative and coordination efforts seem to accumulate over-proportionally (from a certain number of employed professors onwards). In the range of medium-sized BuSs, the research outputs of individual BuS professorships (including respective research assistants) rather add up. This conclusion is plausible in view of the fact that in the German academic discipline of Business Administration, professors usually work individually. Even though some projects are in fact undertaken in cooperation, often this is the case within (their) professorships or between professorships of different universities. Cooperation between different professorships of the same BuS or university are less frequently observed.

Cooperations within a BuS of one university could be promoted, for instance, in terms of joint supervisions of PhD students by multiple professors in postgraduate programs, a concept launched by some universities several years ago. A crucial prerequisite to foster joint research projects between professors within the same BuS is a thematically close profile and focus of a BuS in corresponding research clusters; as initialized by the German federal government within the scope of the "Excellence Initiative". Such clusters could facilitate similarities in content and/or methodology of research. Likewise, this trend towards specialized BuSs and closely cooperating research clusters (or research areas) has been initiated by several universities—in many in conjunction with the introduction of Bachelor and Master programs in Germany. In BuSs, such programmes and convergences of research topics could produce synergistic effects among the participating professorships. Of course, higher coordination requirements arising from administrative processes, for instance in the form of meetings and workshops, could have a negative impact. But the nature and extent of the de facto effects still require additional research. Likewise, the impact of such developments on corresponding properties of returns to scale therefore cannot be assessed yet, posing an interesting starting point for future research.

Further research perspectives arise from the previously discussed limitations of our study. In this paper, we have therefore assumed the data to be deterministic, which

represents a necessary presumption in the application of DEA. A modification could now be made to the effect that SFA, which allows stochastically distributed datasets to be analysed, is used in lieu of DEA. The results of a SFA analysis can then be used to check our findings and, if necessary, to further validate them. The fact that a concrete production function must be assumed ex ante, however, is a disadvantage of using SFA. When DEA is used, by contrast, the orientation of the applied DEA models can be changed. In this paper, we have assumed that only the output can be changed (in the short- to middle-term); for the sensitivity analyses, we also investigated which assertions result from exclusively varying the inputs. Research into scale efficiencies in the event of simultaneous changes to the input and the output, on the other hand, remains to be conducted. In that case, unoriented DEA models must be used to calculate scale efficiencies.

As previously described in the discussion of limitations, the progression of the research production function, and therefore the resulting returns to scale, depend on the input and output factors that are used in the corresponding analyses. For example, qualitative aspects of research could be more closely taken into consideration in future and continuative analyses. In this regard, the third-party funds indicator could be differentiated such that the third-party funds granted by different funders are separated and/or differently weighted because, for example, the German Research Foundation (DFG) sets high standards for scientific rigour in the context of assessing a third-party funding proposal (Joerk and Wambach 2013).

In principle, the CHE has used quantitative measures until now with respect to compiling the indicator set and particularly regarding publication activities, which is to say that published articles are counted (partially quality weighted). Including such indicators in the performance assessment of BuSs could produce disincentives insofar as the research results are divided and published in the smallest possible fragments. The number of publications could therefore be increased correspondingly. This effect could be intensified in the underlying CHE 2007 and 2010 datasets, given that two quantitative indicators are used there to determine publication activities. To reduce these effects, information about the impact or outcome of the publications would be desirable. Referring to this, the citations that a research article receives are a frequently discussed indicator in scientometrics (van Raan 1996). Nosek et al. (2010), for instance, point out that citations represent an impact indicator that is "valid, relatively objective, and, with existing databases and search tools, straightforward to compute." Therefore, it still has to be investigated which impact an inclusion of citations exerts on the properties of returns to scale of BuSs' research production.

As previously outlined, the analysed research productivity of a university generally constitutes only one part of the comprehensive spectrum of academic activities, which furthermore includes teaching and administration. In this sense, the used datasets could be expanded by adding indicators that represent teaching and administration outputs in order to analyse properties of returns to scale with reference to these additional outputs. With respect to teaching, Johnes (2006), for example, uses two indicators consisting of the alumni graduation marks in an offered course of study as output and the students' school marks when applying for a course of study as input. Other qualitative measures, such as evaluation results, dropout rates, supervisory relationships or diversity of the teaching portfolio, would also be conceivable in corresponding analyses. It should be noted, however, that to date there is no recognized set of productivity and performance indicators for teaching, as opposed to research. This is primarily due to the fact that teaching and student learning are tightly interwoven with one another in a co-productive process.

That our results are not random is shown by the inclusion of different consecutive periods on the one hand and on the other hand by multiple robustness and sensitivity analyses. As previously described in the introduction, analyses of returns to scale on

research production processes are subject to particular country- and discipline-specific environmental conditions. If these do not (significantly) deviate from each other, our results can essentially be transferred to other objects of research, for instance other social sciences in Germany, such as economics. However, this does not apply to equipment-intensive disciplines, such as engineering, which, due to funding for machinery or experimental setups, are characterized by considerably higher third-party funding and therefore, where applicable, induced economies of scale and scope. Regarding national specifics, European university systems are comparable with the German one above all others and should therefore exhibit similar returns to scale of research. To that effect, Bonaccorsi and Daraio (2005) also found constant returns to scale for Business Administration research production in Italy. But it is questionable whether our results, for example, are applicable to Business Administration research in the United States of America, where the university system is characterized by a substantial degree of (highly competitive) privatisation. Whether, in particular cases, similar results can be observed for other academic disciplines and/or countries, however, requires further research.

Appendix

See Tables 3, 4 and 5.

Table 3 CHE (2004) Business Administration research dataset for 2001 to 2003

	Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	PhD dissertations (number)	Third-party funds (thousand €)
1	Aachen	24.0	71.00	187.83	29	1902.0
2	Augsburg	42.0	171.00	266.25	49	824.1
3	Bamberg	27.0	66.25	83.50	24	1035.0
4	Bayreuth	30.0	87.00	111.17	38	1535.1
6	**Berlin FU**	**42.0**	**84.00**	**316.33**	**33**	**1307.7**
7	**Berlin HU**	**33.0**	**96.00**	**95.08**	**18**	**2510.7**
8	Berlin TU	29.0	81.00	214.00	46	480.0
9	Bielefeld	18.0	48.00	106.00	10	58.2
12	Bruchsal	18.0	8.00	32.00	1	1595.1
13	Chemnitz	24.0	69.00	183.33	25	2651.1
16	Dortmund	24.0	78.00	85.92	27	686.1
17	Dresden	28.0	83.00	269.08	19	3846.9
18	Duisburg	61.0	143.00	317.75	46	1569.9
19	Düsseldorf	31.0	68.00	153.33	10	177.9
20	Eichstätt-Ingolstadt	50.0	96.00	290.08	44	1227.0
21	**Erlangen-Nürnberg**	**43.0**	**184.50**	**338.17**	**40**	**1811.1**

Table 3 continued

Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	PhD dissertations (number)	Third-party funds (thousand €)
22 Essen	27.0	114.00	180.75	19	2658.9
23 Flensburg	18.0	15.00	31.83	9	1400.1
25 Frankfurt Main Uni	**82.0**	**198.00**	**276.00**	**52**	**7598.1**
26 Frankfurt Oder	27.0	67.00	81.50	24	3552.0
27 Freiberg	27.0	43.50	171.00	27	1016.1
28 Giessen	22.0	68.15	154.08	18	111.3
29 Göttingen	30.0	100.50	141.50	54	2161.8
30 Greifswald	29.5	41.25	107.17	14	2067.9
32 Halle-Wittenberg	25.0	64.00	137.92	11	1062.6
34 Hamburg	40.0	177.49	261.00	41	1364.4
38 Hohenheim	35.0	73.00	202.58	39	4304.1
39 Ilmenau	42.0	123.00	110.08	17	2293.5
40 Jena	24.0	74.30	114.67	12	1622.1
43 Kiel	**24.0**	**40.00**	**151.83**	**14**	**338.7**
44 Köln	**51.0**	**201.48**	**311.67**	**71**	**3232.2**
46 Leipzig HH	23.5	63.50	91.50	15	2917.2
47 Leipzig Uni	34.0	53.00	56.00	29	1310.7
48 Magdeburg	29.0	96.00	92.25	14	1299.9
49 Mainz	24.0	69.00	131.08	36	863.1
50 Mannheim	**48.0**	**254.50**	**476.58**	**75**	**6579.0**
51 Marburg	31.0	61.01	144.25	29	1131.6
52 München LMU	**52.0**	**273.00**	**432.17**	**77**	**4090.2**
53 München TU	36.0	106.00	384.33	38	9108.3
55 Münster	**39.0**	**144.00**	**417.25**	**75**	**1928.1**
56 Oestrich-Winkel	51.0	142.00	351.92	59	2988.0
57 Oldenburg	43.0	64.00	62.83	37	7029.6
58 Osnabrück	24.0	45.92	78.92	8	1437.3
59 Paderborn	24.5	42.75	175.42	18	1421.1
60 Passau	32.0	106.10	99.67	22	2163.9
61 Potsdam	18.0	39.00	137.17	16	291.0
62 Regensburg	27.0	81.00	148.83	40	1002.9
63 Rostock	27.0	74.25	182.00	23	315.3
64 Saarbrücken	51.5	183.50	372.75	50	6695.1
65 Siegen	49.0	44.75	191.00	9	913.5
66 Stuttgart	22.0	51.00	269.92	33	2182.2
67 Trier	27.0	76.50	151.00	24	4224.9

Table 3 continued

Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	PhD dissertations (number)	Third-party funds (thousand €)
68 Tübingen	24.0	61.00	116.33	18	803.1
70 Vallendar	**43.0**	**142.00**	**654.83**	**60**	**876.6**
73 Würzburg	18.0	55.50	96.83	24	1552.8

All (total) values embody the 3-year sum of respective absolute indicator values over the incorporated periods of 2001–2003

Highlighted in bold: Top 10 most reputable BuSs

FU Free University, *HH* Commercial College, *HU* Humboldt University, *LMU* Ludwig Maximilian University, *TU* University of Technology

Table 4 CHE (2007) Business Administration research dataset for 2004 to 2006

Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	Internationally visible publications (number)	PhD dissertations (number)	Third-party funds (thousand €)
1 Aachen	36.00	110.00	241	16	49	3239.0
2 Augsburg	40.00	168.00	299	38	55	4048.4
3 Bamberg	26.00	62.50	145	7	25	487.7
4 Bayreuth	29.00	105.50	184	22	29	1139.5
6 Berlin FU	**47.80**	**91.20**	**274**	**22**	**34**	**2299.6**
7 Berlin HU	**45.00**	**52.62**	**180**	**26**	**21**	**2363.5**
8 Berlin TU	27.00	93.00	188	52	54	5685.0
10 Bochum	30.00	118.00	237	4	32	1508.0
11 Bremen	34.00	53.50	231	6	43	1901.0
13 Chemnitz	27.00	73.65	154	7	38	2815.1
14 Clausthal	20.00	33.00	67	4	5	283.5
16 Dortmund	24.00	78.00	259	4	39	1914.0
17 Dresden	33.00	83.75	162	5	24	2592.2
18 Duisburg	51.50	127.99	281	11	33	1962.5
19 Düsseldorf	21.00	48.00	70	2	22	280.8
20 Eichstätt-Ingolstadt	51.00	152.00	222	7	48	592.4
21 Erlangen-Nürnberg	36.00	140.50	320	23	106	4095.6
22 Essen	27.00	97.75	122	42	25	9188.1
23 Flensburg	31.00	13.00	93	6	14	1669.4
25 Frankfurt Main Uni	**80.00**	**181.00**	**391**	**49**	**86**	**8268.7**
26 Frankfurt Oder	28.00	71.00	105	10	35	2255.3
27 Freiberg	18.00	38.00	100	2	17	896.7
28 Gießen	26.00	65.83	192	3	22	418.8

Table 4 continued

Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	Internationally visible publications (number)	PhD dissertations (number)	Third-party funds (thousand €)
29 Göttingen	30.00	128.32	198	11	60	1955.4
30 Greifswald	28.00	43.00	101	14	9	1290.0
31 Hagen	38.00	166.15	307	21	20	539.1
32 Halle-Wittenberg	25.00	70.00	79	3	27	197.5
33 Hamburg UBW	90.00	150.00	78	2	45	837.7
35 Hamburg Uni Wi/Pol	38.00	26.00	67	4	24	930.4
36 Hamburg Uni WiWi	49.50	167.00	409	31	47	2400.9
38 Hohenheim	41.00	81.25	332	4	44	4142.0
39 Ilmenau	44.00	127.00	183	9	12	850.2
40 Jena	24.50	66.83	199	7	7	725.7
41 Kaiserslautern	30.00	100.00	182	10	21	3512.3
43 Kiel	**28.50**	**62.25**	**202**	**19**	**20**	**850.0**
44 Köln	**60.00**	**263.85**	**525**	**24**	**79**	**4078.0**
46 Leipzig HH	32.00	64.00	140	1	27	316.4
47 Leipzig	42.00	58.50	142	6	24	627.5
48 Magdeburg	38.50	105.00	225	15	14	1510.0
49 Mainz	29.00	86.50	115	4	24	1067.9
50 Mannheim	**67.00**	**237.00**	**681**	**38**	**88**	**6134.0**
51 Marburg	35.00	51.00	161	2	20	610.0
52 München LMU	**56.00**	**247.50**	**425**	**20**	**105**	**3366.9**
53 München TU	**59.00**	**164.12**	**364**	**22**	**79**	**15,909.8**
54 München UBW	54.00	70.50	90	1	43	240.5
55 Münster	**47.00**	**148.50**	**482**	**21**	**98**	**4036.7**
56 Oestrich-Winkel	100.00	156.00	433	15	117	16,615.9
57 Oldenburg	19.00	39.75	198	5	33	2685.6
58 Osnabrück	30.00	41.95	31	3	9	923.0
59 Paderborn	43.00	72.25	128	9	47	1197.4
60 Passau	23.00	64.00	161	11	18	1513.3
61 Potsdam	20.00	35.60	108	8	38	1233.5
62 Regensburg	33.00	94.50	209	10	43	1231.8
63 Rostock	22.92	47.82	106	3	52	1223.4
64 Saarbrücken	48.50	175.00	546	20	37	2847.5
65 Siegen	45.00	48.50	204	15	25	3731.0
66 Stuttgart	20.00	65.60	166	5	28	2023.4
67 Trier	27.00	72.50	94	2	19	1073.0

Table 4 continued

	Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	Internationally visible publications (number)	PhD dissertations (number)	Third-party funds (thousand €)
68	Tübingen	24.00	56.00	113	5	19	586.8
70	**Vallendar**	**44.50**	**164.84**	**615**	**32**	**100**	**2740.3**
71	Witten-Herdecke	29.25	71.71	158	12	57	13,291.0
72	Wuppertal	49.00	85.50	170	6	29	1054.6
73	Würzburg	21.00	50.00	118	3	23	1261.2

All (total) values embody the 3-year sum of respective absolute indicator values over the incorporated periods of 2004–2006. The internationally visible publications were provided by the bibliometric team at Forschungszentrum Jülich

Highlighted in bold: Top 10 most reputable BuSs

FU Free University, *HH* Commercial College, *HU* Humboldt University, *LMU* Ludwig Maximilian University, *TU* University of Technology, *UBW* University of the German Armed Forces, *Uni* University, *Wi/Pol* Economics and Politics, *WiWi* Business Sciences

Table 5 CHE (2010) Business Administration research dataset for 2007 to 2009

	Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	Internationally visible publications (number)	PhD dissertations (number)	Third-party funds (thousand €)
1	**Aachen**	**39.00**	**139.00**	**160.32**	**29**	**66**	**5765.4**
2	Augsburg	40.00	193.00	166.08	36	60	6051.0
3	Bamberg	27.00	71.00	144.51	14	31	632.0
4	Bayreuth	37.00	104.00	92.16	23	41	4152.8
5	Berlin ESCP-EAP	28.00	35.50	140.25	7	15	381.1
6	**Berlin FU**	**48.00**	**79.57**	**212.43**	**28**	**52**	**3309.6**
7	**Berlin HU**	**38.20**	**60.10**	**107.34**	**18**	**28**	**4368.2**
8	Berlin TU	28.00	104.00	180.75	46	106	11,738.0
9	Bielefeld	17.00	36.25	65.43	56	15	876.0
10	Bochum	32.25	111.00	111.93	5	39	2012.1
11	Bremen	49.00	61.50	190.74	32	87	5641.0
13	Chemnitz	27.00	94.65	145.68	19	55	4995.6
14	Clausthal	24.00	44.75	69.99	6	5	579.0
15	Cottbus	56.00	124.00	71.34	5	12	1680.0
16	Dortmund	40.00	98.60	144.93	15	39	2912.0
17	Dresden	42.00	110.50	224.25	24	55	5290.9
18	Duisburg	48.25	73.15	109.26	96	36	3052.0
19	Düsseldorf	30.00	69.00	203.82	2	27	1145.8
21	Erlangen-Nürnberg	41.00	178.28	277.41	24	96	7326.2

Table 5 continued

Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	Internationally visible publications (number)	PhD dissertations (number)	Third-party funds (thousand €)
23 Flensburg	11.00	7.50	64.74	7	15	1951.5
24 Frankfurt Main SFM	87.00	122.00	272.25	21	12	7223.6
25 Frankfurt Main Uni	91.50	162.25	319.92	86	86	10,847.5
26 Frankfurt Oder	45.00	66.00	193.41	23	26	1989.0
27 Freiberg	30.00	48.00	100.17	7	19	1535.7
28 Gießen	27.00	72.24	99.51	12	21	917.0
29 Göttingen	34.00	117.05	114.18	11	26	1437.7
30 Greifswald	32.00	39.50	90.18	12	14	2249.1
31 Hagen	39.00	165.50	246.00	19	27	552.0
32 Halle-Wittenberg	23.00	69.75	77.67	4	15	296.3
33 Hamburg UBW	57.00	95.00	99.00	8	22	295.0
34 Hamburg Uni	**69.00**	**191.54**	**328.59**	**71**	**64**	**4266.7**
37 Hannover	70.00	183.75	178.74	66	53	3896.3
38 Hohenheim	42.00	141.80	252.09	27	60	2294.0
39 Ilmenau	45.00	116.50	213.51	8	15	596.5
40 Jena	31.00	90.50	214.17	17	21	904.4
41 Kaiserslautern	36.00	88.00	124.50	12	20	4461.0
42 Kassel	48.00	83.38	137.49	10	50	2222.0
43 Kiel	**28.50**	**74.50**	**109.41**	**17**	**21**	**713.9**
45 Konstanz	20.00	17.25	54.18	1	11	786.0
46 Leipzig HH	30.00	96.75	122.67	2	18	2960.6
48 Magdeburg	45.50	106.00	146.49	20	19	2878.5
49 Mainz	58.25	134.08	71.25	16	38	866.3
50 Mannheim	**86.00**	**245.00**	**536.43**	**55**	**123**	**9885.2**
51 Marburg	30.00	52.50	53.49	12	40	663.9
52 München LMU	**49.00**	**251.50**	**479.82**	**269**	**67**	**6705.6**
53 München TU	**27.00**	**93.33**	**263.34**	**24**	**92**	**8545.7**
54 München UBW	33.00	60.00	138.18	7	23	898.2
55 Münster	**49.00**	**157.00**	**394.68**	**35**	**105**	**6168.1**
56 Oestrich-Winkel	54.51	54.51	380.76	28	157	9959.5
57 Oldenburg	57.00	67.45	143.67	21	38	4706.4
58 Osnabrück	27.00	60.00	88.17	5	15	1324.5
59 Paderborn	90.00	153.40	249.18	40	52	4275.0

Table 5 continued

Business School	Professors (staffed positions)	Research assistants (staffed positions)	Nationally visible publications (points)	Internationally visible publications (number)	PhD dissertations (number)	Third-party funds (thousand €)
60 Passau	39.00	80.00	72.66	1	20	760.0
61 Potsdam	22.00	50.85	65.82	18	28	4785.6
62 Regensburg	37.00	110.00	123.99	8	57	2059.8
63 Rostock	28.50	63.50	70.17	11	37	1436.2
64 Saarbrücken	45.00	179.50	318.24	21	45	6003.0
65 Siegen	33.56	48.14	79.68	15	32	1971.0
66 Stuttgart	21.00	95.20	103.74	8	31	1179.5
67 Trier	28.00	73.00	71.34	9	39	1736.9
68 Tübingen	29.00	57.50	83.07	26	24	931.8
69 Ulm	22.00	36.00	100.41	14	20	646.7
70 Vallendar	**78.00**	**113.00**	**510.99**	**30**	**92**	**6903.0**
71 Witten-Herdecke	29.85	50.98	82.17	13	49	2480.0
72 Wuppertal	45.00	53.50	158.16	16	23	1193.5
73 Würzburg	19.00	69.88	122.58	10	20	1759.0

All (total) values embody the 3-year sum of respective absolute indicator values over the incorporated periods of 2007–2009. The internationally visible publications were provided by the bibliometric team at Forschungszentrum Jülich

Highlighted in bold: Top 10 most reputable BuSs

ESCP-EAP Ecole Superieure de Commerce de Paris-Ecole des Affaires de Paris, *FU* Free University, *HH* Commercial College, *HU* Humboldt University, *LMU* Ludwig Maximilian University, *SFM* School of Finance and Management, *TU* University of Technology, *UBW* University of the German Armed Forces, *Uni* University

References

Abbot, M., & Doucouliagos, C. (2003). The efficiency of Australian universities: A Data Envelopment Analysis. *Economics of Education Review, 22*(1), 89–97.

Abramo, G., Cicero, T., & D'Angenlo, C. A. (2011). A field-standardized application of DEA to national-scale research assessment of universities. *Journal of Informetrics, 5*(4), 618–628.

Abramo, G., Cicero, T., & D'Angelo, C. A. (2012). Revisiting size effects in higher education research productivity. *Higher Education, 63*(6), 701–717.

Abramo, G., D'Angelo, C. A., & Di Costa, F. (2014). Investigating returns to scope of research fields in universities. *Higher Education, 68*(1), 69–85.

Ahn, T., Charnes, A., & Cooper, W. W. (1988). Some statistical and DEA evaluations of relative efficiencies of public and private institutions of higher learning. *Socio-Economic Planning Sciences, 22*(6), 259–269.

Ahn, H., Dyckhoff, H., & Gilles, R. (2007). Datenaggregation zur Leistungsbeurteilung durch Ranking: Vergleich der CHE- und DEA-Methodik sowie Ableitung eines Kompromissansatzes. *Zeitschrift für Betriebswirtschaft, 77*(6), 615–643.

Banker, R. D., Charnes, A. C., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science, 30*(9), 1078–1092.

Beasley, J. E. (1990). Comparing university departments. *Omega, 18*(2), 171–181.

Beasley, J. E. (1995). Determining teaching and research efficiencies. *Journal of the Operational Research Society, 46*(4), 441–452.

Berghoff, S., Giebisch, P., Hachmeister, C.-D., Hoffmann-Kobert, B., Hennings, M., & Ziegele, F. (2011). *Vielfältige Exzellenz 2011: Forschung, Anwendungsbezug, Internationalität, Studierendenorientierung im CHE Ranking*. Gütersloh: Centrum für Hochschulentwicklung.

Bonaccorsi, A., & Daraio, C. (2005). Exploring size and agglomeration effects on public research productivity. *Scientometrics, 63*(1), 87–120.

Bonaccorsi, A., Daraio, C., & Simar, L. (2006). Advanced indicators of productivity of universities: An application of robust nonparametric methods to Italian data. *Scientometrics, 66*(2), 389–410.

Bort, S., & Schiller-Merkens, S. (2010). Publish or perish. *Zeitschrift Führung + Organisation, 79*(5), 340–346.

Brandt, T., & Schubert, T. (2013). Is the university model an organizational necessity? Scale and agglomeration effects in science. *Scientometrics, 94*(2), 541–565.

Breu, T. M., & Raab, R. L. (1994). Efficiency and perceived quality of the nation's "TOP 25" national universities and national liberal arts colleges: An application of Data Envelopment Analysis to higher education. *Socio-Economic Planning Sciences, 28*(1), 33–45.

Chalmers, A. F. (1990). *Science and its fabrication*. Buckingham: Open University Press.

Charnes, A. C., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–441.

Clermont, M., & Dirksen, A. (2015). The measurement, evaluation, and publication of performance in higher education: An analysis of the CHE research ranking of business schools in Germany from an accounting perspective. *Public Administration Quarterly*, forthcoming.

Clermont, M., & Dyckhoff, H. (2012a). Erfassung betriebswirtschaftlich relevanter Zeitschriften in Literaturdatenbanken. *Betriebswirtschaftliche Forschung und Praxis, 64*(3), 324–346.

Clermont, M., & Dyckhoff, H. (2012b). Coverage of Business Administration literature in Google Scholar: Analysis and comparison with EconBiz, Scopus and Web of Science. *Bibliometrie – Praxis und Forschung, 1*(1), 5/1–5/19.

Cohen, J. E. (1980). Publication rate as a function of laboratory size in a biomedical research institution. *Scientometrics, 2*(1), 35–52.

Cohen, J. E. (1991). Size, age and productivity of scientific and technical research groups. *Scientometrics, 20*(3), 395–416.

Cook, W. D., & Zhu, J. (2007). Classifying inputs and outputs in Data Envelopment Analysis. *European Journal of Operational Research, 180*(2), 692–699.

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data Envelopment Analysis: A comprehensive text with models, applications, references and DEA-solver software* (2nd ed.). New York: Springer.

Dyckhoff, H., Clermont, M., Dirksen, A., & Mbock, E. (2013). Measuring balanced effectiveness and efficiency of German business schools' research performance. *Zeitschrift für Betriebswirtschaft, Special Issue, 3*(2013), 39–60.

Dyckhoff, H., Rassenhövel, S., & Sandfort, K. (2009). Empirische Produktionsfunktion betriebswirtschaftlicher Forschung: Eine Analyse der Daten des Centrums für Hochschulentwicklung. *Zeitschrift für betriebswirtschaftliche Forschung, 61*(1), 22–56.

Fandel, G. (2007). On the performance of universities in North Rhine-Westphalia, Germany: Government's redistribution of funds judged using DEA efficiency measures. *European Journal of Operational Research, 176*(1), 521–533.

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

Frey, B. S. (2007). Evaluierungen, Evaluierungen … Evaluitis. *Perspektiven der Wirtschaftspolitik, 8*(3), 207–220.

Gilles, R. (2005). *Performance Measurement mittels Data Envelopment Analysis: Theoretisches Grundkonzept und universitäre Forschungsperformance als Anwendungsfall*. Lohmar: Eul.

Gutenberg, E. (1983). *Grundlagen der Betriebswirtschaftslehre. Band I: Die Produktion, 24th edition*. Berlin/Heidelberg: Springer.

Gutierrez, M. (2007). Messung der Effizienz von Professuren mittels Data Envelopment Analysis. *Zeitschrift für Betriebswirtschaft, Special Issue, 5*(2007), 101–129.

Jarwal, S. D., Brion, A. M., & King, M. L. (2009). Measuring research quality using the journal impact factor, citations and 'Ranked journals': Blunt instruments or inspired metric? *Journal of Higher Education Policy & Management, 31*(4), 289–300.

Joerk, C., & Wambach, A. (2013). DFG-Förderung in den Wirtschaftswissenschaften: Fakten und Mythen zur Förderpraxis. *Perspektiven der Wirtschaftspolitik, 14*(1–2), 99–117.

Johnes, J. (2006). Measuring teaching efficiency in higher education: An application of Data Envelopment Analysis to economic graduates from UK universities 1993. *European Journal of Operational Research, 174*(1), 443–456.

Johnes, G., & Johnes, J. (1993). Measuring the research performance of UK economics departments: An application of Data Envelopment Analysis. *Oxford Economic Papers, 45*(2), 332–347.

Johnston, R. (1994). Effects of resource concentration on research performance. *Higher Education, 28*(1), 25–37.

Kao, C., & Hung, H.-T. (2008). Efficiency analysis of university departments: An empirical study. *Omega, 36*(4), 653–664.

Kieser, A. (2012). JOURQUAL: Der Gebrauch, nicht der Missbrauch, ist das Problem. Oder: Warum Wirtschaftsinformatik die beste deutschsprachige betriebswirtschaftliche Zeitschrift ist. *Die Betriebswirtschaft, 72*(1), 93–110.

Kyvik, S. (1995). Are big university departments better than small ones? *Higher Education, 30*(3), 295–304.

Laband, D. N., & Lentz, B. F. (2003). New estimates of scale and scope in higher education. *Southern Economic Journal, 70*(1), 72–183.

Longlong, H., Fengliang, L., & Weifang, M. (2009). Multi-product total cost functions for higher education: The case of Chinese research universities. *Economics of Education Review, 28*(4), 505–511.

Madden, G., Savage, S., & Kemp, S. (1997). Measuring public sector efficiency: A study of economics departments at Australian Universities. *Education Economics, 5*(2), 153–168.

Marginson, S., & van der Welde, M. (2007). To rank or to be ranked: The impact of global rankings in higher education. *Journal of Studies in International Education, 11*(3–4), 206–329.

Meng, W., Zhang, D., Qi, L., & Liu, W. (2008). Two-level DEA approaches in research evaluation. *Omega, 36*(6), 950–957.

Nosek, B. A., Graham, J., Lindner, N. M., Kesebir, S., Hawkins, C., Hahn, C., & Tenney, E. R. (2010). Cumulative and career-stage citation impact of social-personality psychology programs and their members. *Personality and Social Psychology Bulletin, 36*(10), 1283–1300.

Rassenhövel, S. (2010). *Performancemessung im Hochschulbereich: Theoretische Grundlagen und empirische Befunde*. Wiesbaden: Gabler.

Ray, S. C., & Jeon, Y. (2008). Reputation and efficiency: A non-parametric assessment of America's top-rated MBA programs. *European Journal of Operational Research, 189*(1), 245–268.

Schrader, U., & Hennig-Thurau, T. (2009). VHB-Jourqual2: Methods, results, and implications of the German Academic Association for Business Research's journal ranking. *Business Research, 2*(2), 180–204.

Stolz, I., Hendel, D. D., & Horn, A. S. (2010). Ranking of rankings: Benchmarking twenty-five higher education ranking systems in Europe. *Higher Education, 60*(5), 507–528.

Tavenas, F. (2004). *Quality assurance: A reference system for indicators and evaluation procedures*. Brussels: European University Association.

Tomkins, C., & Green, R. (1988). An experiment in the use of Data Envelopment Analysis for evaluating the efficiency of UK university departments of accounting. *Financial Accountability & Management, 4*(2), 147–164.

Usher, A., & Savino, M. (2006). *A world of difference: A global survey of university league tables*. Toronto: Educational Policy Institute.

Van der Wal, R., Fischer, A., Marquiss, M., Redpath, S., & Wanless, S. (2009). Is bigger necessarily better for environmental research? *Scientometrics, 78*(2), 317–322.

Van Raan, A. F. J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics, 36*(3), 397–420.

Von Tunzelmann, N., Ranga, M., Martin, B., & Geuna, A. (2003). *The effects of size on research performance: A SPRU review*. Brighton: Science Policy Research Project.