

**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**

**MARCEL CLERMONT  
ALEXANDER DIRKSEN**  
RWTH Aachen University

Acknowledgement: The project underlying this article was funded by the Federal Ministry for Education and Research (Germany) under grant number 01PW11014. The responsibility for the contents of the article lies with the authors. In addition, the authors would like to thank two anonymous reviewers whose comments helped crafting the article.

**ABSTRACT**

Universities do not tend to systematically record their own benefits, leaving it to independent institutions to summarize and publish in the form of ratings and rankings information that such institutions have acquired by themselves. In Germany, for the academic discipline of Business Administration, the Centre for Higher Education (CHE) is the major publisher of such rankings. The challenges in the context of performance measurement and evaluation on the one hand as well as the publication of achieved performance (in terms of benchmarking) on the other hand show considerable analogies to those of managerial cost and activity accounting as well as external financial reporting. Since an analysis of CHE's potential compliance with the relevant accounting principles has not yet been undertaken, the objective of this paper is to systematically and critically analyze from an accounting perspective the method by which CHE actually evaluates research performance of business schools in Germany. Accordingly, we will demonstrate that standard accounting principles, such as completeness – which is an integral requirement of faithful representation – or consistency, are not continuously satisfied. Moreover, classifications of research cost objects, that are made within the meaning of cost and activity accounting, are not unequivocal and can, under certain circumstances,

result in creating misguided incentives for the participating business schools.

Keywords

Business Schools, Centre for Higher Education, Financial Reporting, Research Ranking, IFRS-Framework, Managerial Accounting, Common Accounting Standards

JEL-classification

I23, L31, M41

### **RESEARCH ISSUE AND LITERATURE GAP**

In the context of the implementation of accounting methods and principles by university managements, the possibilities and limitations of university performance measurements are frequently discussed. Whereas the costs, e.g. in terms of staff assignment or expenses, can be recorded fairly easily, the assessing of a university's benefits is more challenging, because basically no exogenous market prices for such benefits exist. The conduction of performance measurements in practice is usually done by external institutions rather than by the universities themselves. Accordingly, such institutions acquire and aggregate the necessary data under their own name before publishing their findings. Generally adhering to a specific method of evaluation, they aggregate predefined performance criteria in order to create ratings or rankings of whole universities or individual departments.

There are numerous different rating and ranking systems for evaluating and comparing universities both in an overall manner or an academic discipline-specific approach. For instance, the US News and World Report publishes rankings that assess US-American colleges and primarily focus on the quality of education by using indicators that are based on facts and surveys. In contrast, the National Research Council in USA evaluates universities' doctoral programs by academic discipline, using various facts and survey-based indicators. Worldwide attention is paid to the QS World University Ranking and the Academic Ranking of World Universities (Shanghai Ranking),

the former evaluating diverse aspects of performance and the latter primarily addressing research performance.

The stated objective of the evaluating institutions is usually that of creating transparency of a university's performance in order to ensure comparability. These institutions perceive themselves to be purely information mediators for interested stakeholders, which includes prospective students, and government bodies as the main financiers. Even though it is not the primary objective, these evaluations do create pressure on the analyzed universities to perform well in accordance with the criteria used and thus to achieve favorable assessments in subsequent evaluations. Hence, universities and/or departments compete to achieve their objectives. In this context, such university assessments or rankings are becoming vitally important in terms of public recognition and ultimately a university's reputation. A high university-specific or department-specific reputation is certainly useful, especially if the competitive situation makes it necessary to attract qualified students and outstanding researchers, and even for retaining governmental financial support in an era of decreasing state contribution.

Performance and improvement incentives emerging from evaluations are not a problem per se. However, this premises the disclosure of the respective performance criteria and of the procedure used for acquiring and aggregating the data for creating the overall evaluation. It also has to be mandatorily assured that the evaluated universities or departments are, as peer groups, rivals in competitive situations, and therefore comparable with regard to their services and objectives. However, such clarity is non-existent in Germany's higher education sector.

Due to this lack of clarity, such rankings are controversially discussed by the academic community in general and more specifically by the evaluated scientists themselves. Especially the key performance indicator-based measurements of academic research performance are the subject of critique (e.g. Frey, 2007; Jarwal, Brion, & King, 2009; Kieser, 2012). In addition, there are several studies that compare the quality of university ranking procedures (e.g. Tavenas, 2004; Usher &

Savino, 2006; Stolz, Hendel, & Horn, 2010). But, since these studies are not based upon a theoretical foundation, they have rather a practical character in terms of benchmarking or product tests. This is not least due to the fact that to date a theory of ranking development or performance measurement is still missing.

However, the resulting challenges for ranking institutions in the context of performance measurement and evaluation on the one hand as well as the publication of achieved performance on the other hand, show considerable analogies to those of managerial cost and activity accounting as well as of external financial reporting. Within the scope of cost object accounting, this becomes obvious with regard to the usage of different cost allocation principles (i.e. inclusion of direct costs versus indirect costs) and the choice of adequate aggregation methods. Accordingly, an academic discipline-specific classification of university departments would involve problems similar to those of the allocation of either itemized-only or overhead costs to different cost objects. For example, it is unclear as to which professorships should actually be allocated to a particular department: only single-discipline professorships, e.g. that of Finance and Accounting, or also interdisciplinary professorships, e.g. that of Business Informatics? Reports based on these calculations are only relevant and reliable for stakeholders if they comply with common accounting standards, for instance via the consistent application of the laws and options for capitalization. This analogy also affects the publication activities both intertemporally for a single department and intratemporally for a comparison of diverse departments.

Although these examples of the parallels between (internal and external) accounting on the one hand and performance measurements and publication in the context of universities in the form of rankings on the other hand are obvious, the actual relationship has to date not been analyzed. Therefore, we are addressing the following questions:

- Which accounting principles and findings are applicable to university rankings? How could a framework for critically

analyzing university rankings be derived on the basis of these principles and findings?

Because of the heterogeneity of different country-specific and academic discipline-specific ranking procedures, an abstract, generic analysis would be inappropriate. Therefore, in this paper we exemplify the generation and the use of such a framework by investigating the research ranking for German business schools (hereafter BuSs) of the Centre for Higher Education (CHE). This ranking is one of the most important from the perspective of the Business Administration academic discipline in Germany. CHE is a non-profit organization founded by the Bertelsmann Foundation and the foundation that supports the German Rectors' Conference (an association of universities and higher education institutions in Germany). CHE claims to be an independent institution that is not bound by directives. One major aspect of its mission is the development of university rankings. Therefore, CHE acquires and evaluates comprehensive data regarding the teaching performance and research performance of universities' departments on a three-year cycle. The results are published on the CHE's website and in the German magazine *Die Zeit*. Based on a detailed analysis of its dataset, CHE generates a ranking focused exclusively on research. It has been published four times up to now for BuSs in Germany. In principle, the research ranking does not intend to address a specific audience. Rather, it is meant to supply information about the research performance of a BuS for any interested university stakeholder. For these stakeholders, CHE intends to create transparency with regard to university performance especially in terms of profiles and core competencies. Meanwhile, ranking results seem to be used particularly by German governmental bodies and university managements in order to assess research performance between different BuSs in a comparative manner. Based upon these assessments, several governmental institutions and universities derive objectives and performance agreements, thus coopting the results of the CHE research ranking.

Similarly to other rankings, the CHE ranking has generated controversy in the academic community. Besides positive reviews, particularly with reference to the transparent

and multidimensional procedure (e.g. Marginson & van der Welde, 2007; Stolz et al., 2010), the ranking is also regarded critically. In particular, the method that CHE uses to aggregate key performance indicators is under discussion (e.g. Ahn, Dyckhoff, & Gilles, 2007; Dyckhoff, Clermont, Dirksen, & Mbock, 2013). Due to the attention paid to the CHE rankings on the one hand as well as the prevailing critique on the other hand, numerous professional associations have recently recommended that the corresponding departments for the academic disciplines should not participate in the surveys performed to create the rankings until further notice, e.g. the German Sociological Association, the German Chemical Society, the German Society for Educational Science. Moreover, several universities are already boycotting the CHE surveys entirely, e.g. the University of Cologne and the University of Hamburg.

In spite of the critique having resulted in a boycott of the CHE surveys by some professional associations and even entire universities, a comprehensive and structured analysis of the strengths and weaknesses of the CHE ranking methodology has not yet been conducted. Only Rassenhövel (2010) analyzes the method from the perspective of production theory. However, a detailed examination of the CHE's potential compliance with the relevant accounting principles has not yet been investigated in the scientific literature. Hence, we address the following further questions:

- How might the CHE research ranking of BuSs be assessed from an accounting point of view? Which relevant accounting principles are satisfied and which are violated? What improvement suggestions for the CHE ranking procedure and publication could be proffered?

Summing up, the objective of our paper consists of developing an accounting-based framework for systematical analyses of university rankings and, on that basis, analyzing the CHE's research ranking method for German BuSs. We examine this research ranking with regard to whether relevant methods and principles of managerial accounting as well as external fundamentals of financial reporting are respected in the ranking procedure. Although, the basic considerations and structures are specified for an analysis of the CHE research ranking of BuSs,

they are transferable to rankings of other academic disciplines, organizations, and countries.

Our paper is structured as follows: first, we introduce the methodological design of our critical analysis in section 2. On the basis of the described principles and the structure, we analyze the steps of the CHE ranking method with reference to their potential compliance with fundamental principals in sections 3 and 4. In this context, we also highlight some misguided incentives for BuSs when participating in the CHE data acquisition. Section 5 summarizes the results and discusses possible improvements for forthcoming CHE rankings. In addition, we address the transferability of our framework and findings to other university rankings.

### **METHODOLOGICAL DESIGN OF THE CRITICAL ANALYSIS**

Business accounting constitutes the main information system of companies. It reflects the financial consequences of a company's activities and can be seen as the monetary representation of economic processes. Economically relevant information is captured at the transaction level, aggregated and prepared for particular accounting purposes, and forwarded to the appropriate recipients. Corresponding to the intended audience, the information system can be divided into an external financial reporting and an internal managerial accounting system. With respect to financial reporting, different standard accounting practices are applied. They depend on the legal system involved (Anglo-American case law vs. Roman-European code law) as well as the jurisdiction. In addition to the provision of information for those stakeholders not belonging to corporate management, standard accounting practices often pursue further constitutive secondary objectives; in Germany, for instance, creditor protection is of vital importance. In contrast, managerial accounting supplies decision makers with information within a company which is intended to provide data for decision making as well as for planning and controlling activities (cf., e.g., Alexander & Nobes, 2013, p. 4ff.).

When determining the appropriate criteria for CHE's potential compliance with the relevant accounting principles, German jurisdiction and the location of both the CHE and the ranked BuSs serve as a departure point. According to § 242, Sections 1 and 2, of the German Commercial Code (HGB), basically every German merchant is – and in conjunction with §§ 1-7 HGB all resident companies are – obligated to financially report in accordance with the accounting principles of the HGB. However, the HGB is subject to progressive modifications over time, which arise from individual countries' synchronization of accounting principles. In general, due to the flow of goods and the internationalization of funds, European companies are exposed to increasing pressure to provide information-based international financial reporting, in alignment with International Financial Reporting Standards (IFRS). Extrinsically, this is promoted by the information requirements of foreign stakeholders; intrinsically, it is motivated by the creation of additional strategic perspectives for action. Additionally, as of January 1, 2005, § 315a, Section 2, HGB obligates capital market oriented companies to disclose their consolidated financial statements in accordance with IFRS. In summary, IFRS occupy a position of superior and steadily increasing importance. Beyond that, analogous to the CHE research ranking and deviating from the HGB, the primary objective of IFRS consists of delivering essential information to interested stakeholders. This is why IFRS provide adequate criteria for examining CHE's potential compliance with essential accounting principles.

IFRS regulations in addition to the conceptual framework contain individual standards (IAS/IFRS) as well as corresponding interpretations (SIC/IFRIC/IFRSIC). These standards and interpretations constitute a *lex specialis*, which, in conflict situations, is given preference over the *lex generalis* of the conceptual framework.

However, the special case analysis of the CHE research ranking and other university rankings as well are not congruently transferable to business situations. Accordingly, companies' systems of objectives can vary; the dominant objective of each market-based company can, though, certainly be seen as the realizing of monetary profit. Positive profit is a fundamental



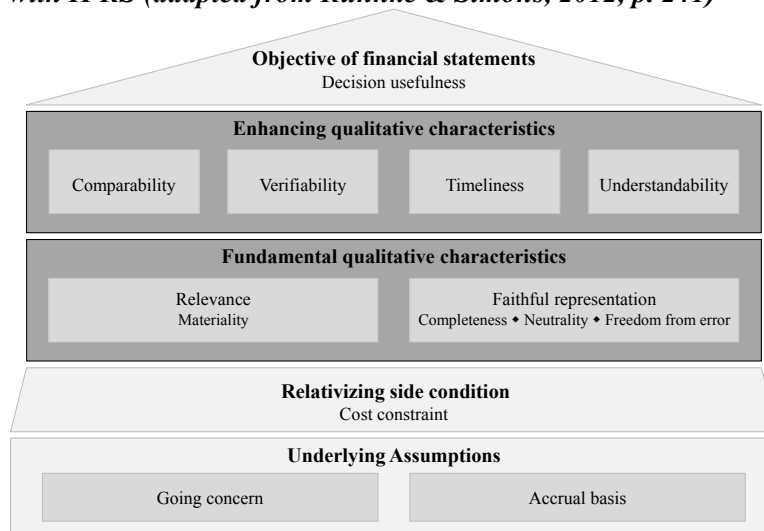
prerequisite to guarantee solvency and consequently ensure the survival, the development, as well as the achievement of further, possibly superior, objectives. In contrast, the deviating intention of non-profit university research is the development of new, publicly available knowledge about the world (Chalmers, 1990, p. 23); financial objectives are only instrumental. Accordingly, reports aligned with IFRS focus on the consequences of financially relevant incidents; the CHE research ranking, in contrast, focuses on the publication of relevant, generally nonmonetary, costs and benefits of research activities.

As a result, the *lex specialis* of IFRS, with its focus on monetary measures, cannot reasonably be applied to an analysis of the CHE research ranking. However, the *lex generalis* codified in the conceptual framework (see International Accounting Standards Board, 2010) defines and describes the theoretical basis of IFRS financial accounting. It is meant to support the application of IFRS regulations to interpret existing standards, to serve as an orientation in the case of the absence of explicit rules, and to develop new standards. In this context, the framework particularly incorporates the fundamental accounting principles, which are to be complied with in IFRS financial statements in order to maximize the information content of reports. Figure 1 shows these principles systematically.

Hence, in analogy to the CHE research ranking, the primary objective of the reporting is the compiling and disclosure of useful information for decision making, especially for controlling and forecasting activities. The objective of stewardship within the framework of both CHE and IFRS carries less weight. Furthermore, the lower part of Figure 1 displays the two basic assumptions of IFRS reporting, namely the going concern as well as accrual basis assumptions. For our analysis, the accrual basis assumption is particularly important because it defines the elementary differentiation concept of the reporting. According to this assumption, costs and – corresponding to the matching principle – benefits are allocated to those temporal periods that they belong to economically, which is not mandatorily the period of the corresponding incoming and outgoing payments. Furthermore, the cost constraint as an elementary relativizing side condition has to be respected; the

usefulness of the disclosure and prior inquiry always ought to exceed the corresponding expenses. In addition, the conceptual framework gives six qualitative characteristics, organized in two hierarchical levels, serving as guidelines to maximize information content.

**Figure 1**  
*A system of fundamental accounting principles in accordance with IFRS (adapted from Ruhnke & Simons, 2012, p. 241)*



These guidelines, highlighted in dark gray in Figure 1, prove to be exceptionally appropriate with reference to an analysis of the CHE research ranking and are therefore used as primary criteria. They are defined as follows:

- **Relevance:** Information should always have confirmatory as well as predictive value. Further, information should be material with regard to the type and quantitative dimension within the scope of the report.
- **Faithful representation:** The provided information should satisfy the requirements of completeness, neutrality, and freedom from error.

Generally, these two fundamental qualitative requirements are in an inverse relationship with each other. The more reliable the representation, the more the corresponding

relevance decreases, and vice versa. Indeed, IFRS postulate a balanced relation between faithful representation and relevance. But in general, particularly because of the fact that only relevant information is to be represented, a certain dominance of relevance over reliability, or more precisely faithful representation, exists. In this respect, it could be stated that relevant information is to be disclosed under the side condition of faithful representation.

These two fundamental qualitative characteristics are underpinned by four supplementary or enhancing qualitative characteristics as attributes which should be sufficiently satisfied:

- Comparability: Reports should be comparable both intertemporally for the same entity and intratemporally for different entities. This primarily finds expression in the principle of consistency. Basically, it addresses identical practices, temporal as well as factual, in similar situations, and it formally concerns presentation and disclosure.
- Verifiability: With reference to transparency, this principle postulates that information must be intersubjectively verifiable – either directly via observation of the reported information or indirectly by checking the calculation procedures.
- Timeliness: This principle addresses the reciprocal nature of relevance and faithful representation. If information is incorporated in reports too late, even if it is generally faithfully represented, its relevance is reduced. In the case of too rapid reporting, under certain circumstances not all essential information is known yet, which leads to a limitation in the faithful representation. In order to keep a balanced relation, it should be considered how to best allow for the information requirements of the stakeholders.
- Understandability: Information should be comprehensible for its audience, taking into account their respective educational background. To achieve this, data should always be clearly, consistently, and concisely classified, declared, characterized, and presented. However, relevant data should not be excluded just because of high complexity and difficult comprehensibility.

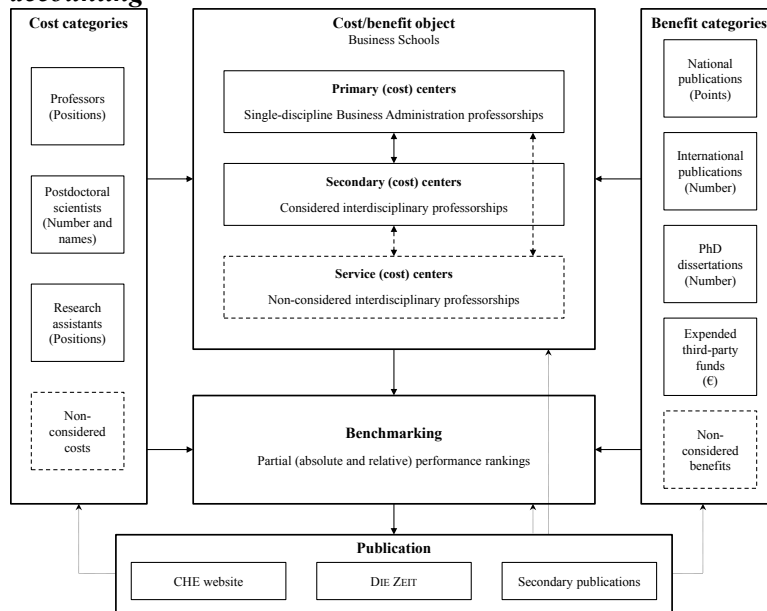
After having derived the relevant criteria from IFRS for evaluating the CHE research ranking, we now have to generate a structural framework for a systematical analysis. Compared to managerial accounting, the procedure for the CHE research ranking shows considerable analogies to the conventional structural accounting system, which consists of cost category, cost center, and cost object accounting, as well as correspondent valuation and allocation problems (cf. Figure 2).

CHE identifies the self-predefined cost and benefit categories via the BuS surveys and own inquiry (on the left and right in Figure 2). In addition to the cost and benefit categories to be considered, the dotted rectangles indicate existing costs and benefits which CHE does not consider to be relevant within the scope of its data acquisition. The considered costs and benefits are generated by the corresponding Business Administration cost centers, i.e. the professorships. The single-discipline Business Administration professorships, the considered interdisciplinary professorships as well as the non-considered interdisciplinary professorships (dotted rectangle) can be distinguished. The double arrows connecting the professorships indicate that they do not conduct research in isolation of each other. Rather, they are integrated into the BuS environment and cooperate in various forms; in the case of interdisciplinary professorships, they are perhaps also involved in research in the department of the other academic discipline. The sum of all Business Administration professorships constitutes the overall respected cost object considered by CHE, i.e. the considered BuS. Using a specific aggregation and evaluation benchmarking process, CHE subsequently creates different absolute and relative performance partial rankings (hereafter PPRs) to derive the individual research profiles and the final research strength of a BuS.

The resulting research rankings are published on the CHE website and in the German magazine *Die Zeit*. Beyond that, the data and ranking results are analyzed and used in secondary literature. The dotted feedback arrows indicate potential incentives stemming from the publications for both the CHE, in alignment with the conception and acquisition of cost and benefit categories as well as the creation of benchmarking and, for the participating BuS, with regard to its own

classification and reported data. In this respect, the CHE method consists of four comprehensive components: category accounting, center accounting and object accounting, and benchmarking. Under consideration of IFRS principles, we will explain and critically analyze these components in sections 3 and 4. Our analyses will focus on the recent research ranking method for BuSs in 2011.<sup>1</sup>

**Figure 2**  
*Structure of the CHE research ranking cost and activity accounting*



**COST AND BENEFIT CATEGORY ACCOUNTING  
IN THE CHE RESEARCH RANKING**

***Definition and Acquisition of Cost and Benefit Categories***

As already shown in Figure 2, CHE considers four benefit and three cost categories in its ranking for the research performance evaluation of a BuS. The figures are gained over a three-year period, in the year preceding the publication of the ranking, i.e. for the 2011 ranking: 2007-2009. The national

publications are identified through a publication analysis of the German language literature database WISO. To execute this bibliometric analysis, CHE requests a BuS to forward the names of its postdoctoral scientists (i.e. including professors) who were engaged in Business Administration research up to a specified date, i.e. for the 2011 ranking: May 31, 2010. Afterwards, the identified articles are converted into publication points by incorporating the number of authors, quantity of pages, and type of article as well as, in the case of journal articles, the quality of the publication medium. The assessment of such quality is based on the journal ranking JOURQUAL2 of the German Academic Association for Business Research (Schrader & Hennig-Thurau, 2009). Supplementarily, the international publication activity of a BuS is indicated by the number of international publications. These are identified by conducting a query on the US-American literature database Web of Science (WoS) on the basis of BuS name. Deviating from the determination of national publications, the resulting number of international publications is not converted into weighted publication points.

The number of PhD dissertations, the amount of expended third-party funds in the form of research grants and contracts, and the cost categories are acquired by surveying a BuS. CHE gathers the PhD dissertations per semester over the six semesters of the survey period. The expended third-party funds are requested separately by type and funder. Within the framework of the identification of the cost categories, on the one hand, CHE records the number and names of the postdoctoral scientists researching in a BuS up to the due date stated above. On the other hand, it records the staffed positions of BuS professors and research assistants as full-time equivalents that are financed by basic funding; both to be reported on December 31 of the survey years. According to Berghoff et al. (2011), the comprehensive method is assisted by an expert advisory committee.

#### ***Critical Analysis of the Cost and Benefit Category Accounting***

From the perspective of decision theory, prior to the conceptually designing of cost and benefit categories, a mandatory prerequisite must be a well-founded and systematical

understanding of the underlying objectives of a BuS. This is essential for cost and benefit categories that are not measurable with a generally accepted monetary denominator. In this context, Ewert and Wagenhofer (2014, p. 32ff.) denote these criteria as “costs and benefits I”; Dyckhoff (2006, p. 126f.) refers to such criteria by using the terms “real costs” and “real benefits”. However, such a well-founded analysis of the essential objectives of a BuS has only been conducted by CHE via consultation with the expert advisory committee. The main objective of scientific research – the generation, publication, and utilization of new public knowledge about the world – is, in contrast with the monetary-focused perception of companies, only measurable indirectly and can only be operationalized by using diverse cost and benefit categories. In this respect, the CHE category set is to be assessed with regard to relevance as well as faithful representation, in the form of freedom from error, completeness, and plausibility of the data acquisition in relation to every single considered cost and benefit category. Moreover, the completeness of the overall category set must be verified in general.

First of all, we need to clarify whether the cost and benefit categories considered by CHE are appropriate for measuring the comprehensive research performance of a BuS, within the meaning of the superior constitutive objective in line with Chalmers (1990). In this regard, we can definitely state that dissemination and discussion of new knowledge within the academic discipline of Business Administration mainly takes place via written articles, particularly journal articles (Bort & Schiller-Merkens, 2010, p. 340). Consequently, due to the predominant importance of BuS publications, the assessment by two benefit categories – national and international publications – appears to be appropriate. Likewise, the implementation of quality weights, at least for national publications, depending on the reputation of the publication medium seems to be expedient.

Concerning the prerequisite of a faithful representation, we further have to scrutinize to what extent particularly completeness and freedom from error are assured in identifying publications through use of the literature databases noted above. In this context, the coverage rate of relevant Business

Administration journal articles in the WoS and WISO databases has been investigated by Clermont and Schmitz (2008) as well as Clermont and Dyckhoff (2012). They analyze which JOURQUAL1 and JOURQUAL2 ranked journals are indexed in these databases. Both studies demonstrate that numerous journals which are considered to be relevant by the members of the German Academic Association for Business Research are not actually covered. Furthermore, the authors find that the coverage rates vary among the different sub-disciplines of Business Administration. Hence, their research results indicate the unfeasibility of a full identification of the publication performance of a German BuS via the databases utilized. In particular, it is conceivable that those scientists and BuSs are advantaged that research and publish primarily in those sub-disciplines which have a high database coverage. CHE seems to be well aware of this problem, since it even states that it requests samples of comprehensive publication activities (Berghoff et al., 2011, p. B4). However, it is questionable as to whether such drawn samples are actually appropriate for representing the publication performance of a BuS and whether they might incorporate unsystematic biases.

Clermont and Dyckhoff (2012) suggest that a combined bibliometric analysis using the WoS and WISO databases would appear to be advantageous. If such an analysis were conducted, duplications would have to be eliminated because both databases index a high number of the same English language journals. However, as CHE does not eliminate duplications, it can be assumed that BuSs publishing in those journals that are indexed in both databases will be unfairly advantaged. This overrepresentation counteracts a faithful representation within the meaning of being neutral and error-free. Furthermore, it possibly restricts the comparability of BuSs. Analogously, single bookings of expenses and multiple bookings of the correspondent incomes are (taking the matching principal into account) naturally inconceivable in a double-entry bookkeeping system.

The last point further illustrates that the terminology used by CHE, that is, “national” and “international” publications, could at least intuitively result in misguided implications. The



first benefit category suggests that only German language articles have been considered or at least articles that have been published in Germany, and that the second benefit category is only taking international publications into account. As pointed out above, this is not the case. Moreover, contradictorily to the CHE terminology, national publications are actually denoted as weighted publication points. Hence, within the meaning of clear understandability, “visible national publication points” and “visible international publications” would be more precise terms for the underlying concepts.<sup>2</sup>

National publications are queried only on the basis of the names of the postdoctoral scientists employed in the field of Business Administration at the respective BuS.<sup>3</sup> Accordingly, any publication points generated by non-PhD research assistants are not covered. This, in turn, implies an incentive for a BuS to appoint at least one postdoctoral scientist as a co-author, regardless of whether he or she has actually contributed to the article.

While for the 2011 ranking, publications are covered from January 1, 2007 to December 31, 2009, the names of postdoctoral scientists working in a BuS are reported considerably later, on May 31, 2010. However, during the time period from January 1, 2010 to May 31, 2010, a BuS will potentially have recruited or released several scientists. Therefore, it is not comprehensible as to why CHE did not select December 31, 2009 as the designated reporting date.

With reference to international publications, the bibliometric approach has been modified. CHE now uses the name of a BuS for publication queries on WoS instead of the names of the postdoctoral scientists. By doing so, the discussed critique regarding the limited selection of scientists considered when identifying essential publications is to be put into perspective. But, since the associated BuS cannot always be explicitly identified from the details provided in literature databases, the database user requires substantial knowledge of the academic discipline of Business Administration in Germany. Otherwise, for instance, non-BuS publications originating from other academic disciplines within the same university might be taken into account. However, if a query on the basis of a BuS

name is preferred to a query on the basis of a scientist's name, it is inconceivable as to why this does not apply for national publications, too.

Likewise, the differing evaluation or rather weighting of the two publication benefit categories is somewhat less justifiable. Thus, international publications are not weighted, whereas the national publications are, by means of diverse criteria. The lack of weighting with regard to international publications suggests that collaborative articles by scientists from the same BuS are implicitly awarded lower values than single-authored or co-authored publications in collaboration with people from outside the BuS (cf. also Rassenhövel, 2010, p. 189). In general, the weightings appear to be reasonable, for instance in order to assess articles by many co-authors differently from those generated in single authorship and in the latter case, it is likely that considerably more time resources were utilized by the correspondent researcher. The time needed to create scientific articles may also have been one major reason for weighting scientific papers using the number of pages involved. But it is questionable as to how these weights can be determined expediently without the existence of generally accepted, standardized pages. The average number of words on a page differs substantially depending on the respective publication medium. Moreover, it is arguable as to how scientific journals should be rated objectively with respect to the quality of the papers published in them. Neither the rating of journals based on their readership opinions – such as in JOURQUAL – nor the rating via impact factors are undisputed (cf., e.g., Boor, 1982; Reedijk & Moed, 2008; Eisend, 2011; Kieser, 2012). Regarding the latter rating method, scientific studies state that in journals that feature high impact factors, numerous articles are hardly cited or even not at all. In fact, only a few frequently cited publications manage to generate high journal impact factors (Baum, 2011).

PhD dissertations per se comply with the objective of educating and training young scientists. The main objective of any doctoral student, however, is to generate new scientific findings and to present these to the public and the academic community through a published dissertation. Therefore, the

number of PhD dissertations can indicate certain achievements of objectives with regard to the research activities of a BuS (Dyckhoff, Rassenhövel, & Sandfort, 2009, p. 48). Nevertheless, the number of PhD dissertations may not be independent from the number of research assistants, who are either financed by (typically governmental) basic funding or third-party funds. Therefore, the resulting significant correlation coefficient of 0.84 at the 1% level between the amount of expended third-party funds and the number of PhD dissertations is intuitively plausible. However, high correlation coefficients of costs and benefits do not necessarily justify their exclusion. Analogously, balance sheet items in accounting are interdependently associated with each other. Thus, in the context of business activities, total revenue and marginal returns (or profits), for example, are almost certain to exhibit high correlations. The core issue is rather whether the data convey additional decision-relevant information regarding the achievement of one or more sub-objectives of research. From our point of view, this applies to PhD dissertations, because realistically, not every research assistant will graduate with a PhD degree. Potential for PhD graduation is dependent on the personality, intellect, and diligence of a research assistant as well as on the support of a professor or supervisor. The record, however, is incomplete in as far as the quality of PhD dissertations cannot be detected simply by counting how many of them there are. As an example: in Germany, it recently became evident that several PhD dissertations from earlier decades did not comply with generally accepted standards. After a thorough assessment by the Examination Committees of the relevant universities, the doctoral degrees were revoked. Only measuring the number of PhD dissertations might intensify such problems of quality, because potentially misguided incentives to lower the requirements of PhD dissertations may be created in order to enhance the value of this benefit category.

With respect to the BuS survey, BuSs are requested to report the number of their PhD graduates. However, the term “PhD graduate” is not precisely defined: the dissertation might have been completed and evaluated, or the candidate might have completed the dissertation and have also taken part in the oral

disputation, or yet another candidate may have completed both stages and already have her or his certificate after publishing the dissertation. A random survey of different BuS employees shows that they do not concur on the respective time stage that CHE is addressing with regard to a “PhD graduate”. Assuming that different BuSs construe this CHE directive in divergent ways, a comparison of these numbers would be restricted. Analogous to valuation options in financial reporting with regard to the allocation of substantial overhead costs to cost objects, varying but consistent BuS practices are irrelevant for individual BuSs over the total period because the sum of all PhDs over all particular periods always equals the amount over the total period, presuming that the PhD student has had her/his dissertation published. The problem arises in the case of varying perceptions within the same BuS over consecutive periods. Then, multiple counts of one PhD dissertation are, in the course of time, possible, and the principle of consistency is violated.

The classification of expended third-party funds as a benefit category is not clear (Hornbostel, 2001; Jansen, Wald, Franke, Schmoch, & Schubert, 2007; Dyckhoff et al., 2009, p. 49). Third-party funds can either constitute a resource effort and therefore a cost category, or they can be construed as a desired research achievement – associated with the raising of funds or predictions of their use – and thus as a benefit category. Hence, it is subject to the respective decision making context or the underlying objectives as to whether third-party funds are to be regarded as costs or benefits.<sup>4</sup> Due to the positive perception of governmental research policy in general, as well as of university administrations in particular, the consideration of third-party funds as a desired benefit category would appear to be satisfactory. Admittedly, in using expended third-party funds as a proxy variable, it is implicitly assumed that the underlying research projects will actually be successful. Accordingly, the incorporation of a yet to be defined measure of project success would be desirable. The funding body and the associated rigor of the scientific review process could potentially induce the nature of a research project. To date, the expended third-party funds are requested separately, but are ultimately disclosed in an aggregated form. Taking into account that third-party funds

issued by funding bodies such as the German Research Foundation or the German Federal Ministry for Education and Research are granted on the basis of a scientific review process (with reference to the former, cf. Joerk & Wambach, 2013), whereas third-party funds issued by the private sector are not necessarily subject to reviews, a distinction between the grants of the diverse funding bodies is advisable. Additionally, the definition of the respective benefit category exhibits certain manipulation opportunities. In this context, CHE does not address the requested third-party funds resulting from consulting services and endowed professorships as research funds, and accordingly deducts these from the total amount of all third-party funds. Irrespective of this deduction, the funds might alternatively be posted in different sub-categories by a BuS.

The cost categories considered by CHE are pure measures of staff assignment in terms of professors and research assistants. There is no doubt that these human resources constitute the primary cost category of Business Administration research. However, the conjoint usage of cumulative and periodic values in the form of number and names of postdoctoral scientists on a specific date as well as staffed positions over a length of time should also be noted. Shifting to financial accounting would equal a kind of hybrid report that combines measures of flow and stock and lead to respective transparency limitations for the stakeholders.

Having extensively discussed the CHE's category set as well as the acquisition of corresponding values, we still have to question the extent to which this underlying set can be regarded as complete within the meaning of a faithful representation of research performance. There are numerous empirical studies that use one criterion or multiple criteria of the CHE ranking. But a detailed, theoretical-conceptual analysis of expedient cost and benefit categories in order to comprehensively measure research performance has been executed in only a few studies. Against this background, Dyckhoff, Rassenhövel, Gilles, and Schmitz (2005) address this issue and also refer to Chalmers's research objective definition. Based on the construct of purpose-rational acting by Weber, Handerson, and Parsons (1964), the authors

derive systematic operational objectives of university research and provide potential indicators (Table 1).

**Table 1**  
*Structure of research objectives and indicators according to Dyckhoff et al. (2005)<sup>5</sup>*

Generating new, generally public knowledge about the world			
Achieving objectives	Using resources economically	Achieving desirable side-effects	Avoiding undesirable side-effects
<b>Publications → Citations</b> Journal articles Monographs Articles in proceedings Working papers Editorships	<b>Staff</b> Professors Research assistants Non-research assistants  <b>Physical resources</b>	<b>Education of young</b> PhDs Habilitations <sup>5</sup>  <b>Enhancement of academic teaching</b>	<b>Unintended consequences</b> e.g. Nuclear energy Genetic engineering Environmental pollution
<b>Acquired third-party funds</b>	<b>Financial resources</b> Basic funds Expended third-party funds	<b>Development of new products and technologies</b> Patents Licences  <b>Transfer of technology</b>	
<b>Presentations</b>		<b>Promotion of international cooperation</b>	
<b>Research awards</b>			

CHE does not use all potentially conceivable cost and benefit categories shown in Table 1, which implies that, in principle, it provides an incomplete picture of the research performance of German BuSs. However, it is questionable as to what extent these cost and benefit categories can improve decision making, at least if CHE ultimately aims to achieve an evaluation and comparison of research performance between BuSs. For example, patents and licenses probably play a significantly smaller role in Business Administration research than they do in Engineering Science. It is rather unlikely that valid findings for a BuS comparison would actually be achieved by analyzing benefit categories of this kind. Physical resources, such as machinery and laboratories, currently exert a minor impact on the research performance of a BuS. However, new methodological approaches in Business Administration research are on the increase, which require considerably more physical resources. For example, this applies to neuropsychological and experimental research approaches, where it might be useful to consider additional cost categories in data acquisition. Other

benefit categories, such as the transfer of technology or the promotion of international cooperation, are rather qualitative in nature and therefore insufficiently quantitatively measurable. In addition, in order to select an adequate dimension of reporting activities and consequently the extent of cost and benefit categories to be acquired, the principle of materiality – in our case for the evaluation of Business Administration research – in conjunction with the cost constraint, must be taken into account. In this respect, for instance, an additional acquisition of the transfer of technologies presumes a transformation into a measurable indicator. Ultimately, it has to be ensured that the costs of data acquisition for additional indicators do not exceed their material impact on the final information content of the ranking. Hence, any extensions of the indicator set should always be undertaken carefully.

#### **COST CENTER AND COST OBJECT ACCOUNTING IN THE CHE RANKING**

##### ***Classification of Business Administration Professorships, Calculations of Performance Criteria, and Derived Benchmarking***

The cost and benefit categories presented and discussed in section 3.1 are assigned to the Business Administration professorships by using an implicit cost center and benefit center accounting process. A direct consideration of a cost center “department” is not possible, since departments or rather BuSs are organized differently. For example, chairs of Mathematical Economics or Business Law could organizationally be allocated to a BuS, although such chairs do not – in a strict sense – belong to the original discipline of Business Administration. In the CHE ranking, the classification of chairs constituting a BuS stems – by self-reporting – from the BuS itself. Thus, each BuS is responsible for the consideration or non-consideration of interdisciplinary professorships, as in the sample case of Business Information Systems. The research costs and benefits of each professorship (center) are allocated to the respective BuS on the basis of cost object and benefit object accounting. Furthermore, CHE utilizes a specific aggregation method for its

benchmarking, which results in different PPRs. These PPRs are the departure point for deriving the research profile of a BuS and for determining that school's ultimate research strength.

The first four of the eight PPRs are generated by taking the four benefit categories into account. For each benefit category and BuS, the arithmetic mean over the three years is calculated. On this basis, four PPRs are derived that represent absolute or effectiveness dimensions. In order to respect the size of an evaluated BuS as well, four additional relative performance criteria are computed. These exhibit relative or efficiency dimensions, corresponding to the absolute benefit categories: national publication points per postdoctoral scientist and year, international publications per scientist (= professors plus research assistants) for three years, the amount of expended third-party funds per scientist and year as well as the PhD dissertations supervised per professor and year. Based on these four relative performance criteria, the respective PPRs are generated. Due to the comparative evaluation of a BuS within benchmarking, CHE classifies a BuS as being "Strong in research" if it belongs to the respective top group in at least half of the eight PPRs. Regarding the four absolute PPRs, the respective top group for each criterion consists of those BuSs that display the highest values. Additionally, their values cumulatively have to account for 50% of the sum of all the values. The other BuSs are divided into a middle and lowest group. The lowest group consists of those BuSs that represent the lowest values and cumulatively account for 10% of the total sum of all values at most. The remaining BuSs belong to the middle group. In reference to the relative performance criteria, also three groups are distinguished: the top group consists of the 25% best BuSs featuring the highest ratios. The middle group represents the next best 50%, and the remaining 25% belong to the lowest group. Beyond this basic procedure, CHE seeks to identify distinct group transitions (CHE, 2014). Therefore, significant differences in the data structure are taken into account by slight deviation from the basic procedure and by adjusting transitions between groups downwardly. BuSs that exhibit similar values with regard to a criterion will be assigned to the



same (higher) group and thus represent the same research strength relative to the corresponding PPR.

***Critical Analysis of Cost Center/Cost Object Accounting and Benchmarking***

The classification of Business Administration professorships, their subsequent allocation to a BuS as well as the aggregation method for benchmarking must be examined critically with reference to relevance, consistency, transparency, and verifiability of the procedure. Hence, allowing discretion when Business Administration professorships are being allocated to a BuS can result in misguided incentives. For example, those BuSs with only a small number of scientists employed are disadvantaged with regard to absolute performance criteria and the corresponding PPRs per se. They might have the incentive to focus on the relative PPRs and to report only those professors who are performing favorably in comparison to the others, in terms of the considered CHE benefit categories.

Since CHE does not identify and publish the names of professors from the evaluated BuSs, it is not possible to conclude which interdisciplinary professorships (and associated research assistants, postdoctoral scientists, and respective benefits) have been considered or not. Also, professors and postdoctoral scientists who are strong in publishing scientific articles could have been reported exclusively by their names and not by their positions. By using these hybrid reports, a BuS could profit from the resulting strength in publication without negatively impacting the respective relative PPRs. In this context, we examined the reported number of scientists over time by BuS. The results show high variations with regard to the analyzed number of scientists in the course of the three previous research rankings. Table 2 illustrates this pattern by means of four BuSs. It remains unclear as to whether these fluctuations are a result of intentional strategic reporting considerations, of unintentionally incorrect datasets, or of modifications in the organizational structure of a BuS. However, due to the smoothing effect of three-year averages, these fluctuations appear to be rather less logically justifiable. They directly influence the consistency of data acquisition in a negative way and thus indirectly affect the

fundamental qualitative characteristic of a faithful representation.

Due to the lack of regulations regarding the classification of BuS professorships and associated scientists, it is de facto impossible to verify the correctness of chosen classifications as well as the number of PhD dissertations and the amount of third-party funds. Referring to financial accounting, this issue is comparable to the allocation of overhead costs in cost object accounting.<sup>6</sup> Such asymmetrically distributed information could, in turn, encourage a BuS to define the costs in terms of scientists restrictively, whereas the benefits of a BuS could be identified less restrictively. This would result in the respective benefits being overstated. Due to the query of international publications on the basis of BuS name, an additional benefit category not reported by the BuS is affected, too, because publications by non-reported scientists could certainly be included in this performance criterion.

**Table 2**  
*Fluctuations in numbers of scientists of four selected BuSs<sup>7</sup>*

	BuS 1		BuS 2		BuS 3		BuS 4	
	Number of scientists	Delta in %	Number of scientists	Delta in %	Number of scientists	Delta in %	Number of scientists	Delta in %
2005	32.8	-	11.0	-	64.3	-	35.7	-
2008	17.3	-47%	22.1	100%	167.3	160%	19.3	-46%
2011	32.9	90%	6.2	-72%	36.3	-78%	41.5	116%

With respect to the subsequent assessment of research performance in the form of benchmarking, it is to determine whether the BuSs fulfill the requirement of comparability and whether the resulting information is presented adequately. First, it should be emphasized that the PPRs taken into consideration by CHE always present a limited view of the total academic performance of a BuS by focusing on research performance. In order to present an enhanced overall view of a BuS, teaching and administration activities would have to be taken into account as well. The underlying cost categories, however, are allocated to the entire spectrum of university activities.<sup>8</sup> In a research evaluation case, this aspect is largely irrelevant if the additional requirements on all the compared BuSs are identical. By

focusing on German BuSs, certain similarities with regard to organizational structures and research objectives may be assumed, but it still remains questionable, for example, whether the challenges and the pressure of other activities at private universities equal those of public universities. If challenges and other required activities are not similar, the validity of research rankings is going to be limited as a matter of course.<sup>9</sup>

Within the benchmarking conducted, all considered performance criteria are equally weighted. To achieve the designation “Strong in research”, this implies that a BuS has to perform well in diverse (at least 50%) absolute and relative performance criteria rather than having a specialized research profile. But it is quite conceivable that a BuS will develop individual research orientations and focuses, leading to diverse research objectives between different BuSs. These specific research objectives and resulting core competencies can, in turn, lead to heterogeneous basic parameters that impact on research performance.<sup>10</sup>

Moreover, teaching objectives and profiles may influence research performance in different ways, negatively or positively, and thus limit a comparison between different BuSs. For example, due to their typically higher number of students, the BuSs of public universities generally dedicate higher teaching workloads than those of private universities. This, in turn, reduces the remaining time assigned to research, and could possibly result in a weaker research performance. However, large numbers of students as well as interdisciplinary integrations of study programs could also create positive side effects for research activities, in terms of economies of scale and scope. For instance, conjoint research potential is recognized and capitalized on in interdisciplinary research collaborations within universities. In addition, the non-university regional environment and its economic strength can decisively impact the research performance (Rassenhövel, 2010, p. 160f.).

The use of elementary ratios within the relative PPRs in order to link together and compare the productivities of differently sized BuSs is intuitively understandable and comprehensible. However, this implicitly assumes constant returns to scale of the measured processes (Ahn et al., 2007, p.

623). In general, there are different plausible reasons for any kind of returns to scale: e.g. economies of scale and scope (promoting task-sharing and cooperation opportunities) for increasing and disproportionately escalating coordination activities for decreasing returns to scale. Empirical studies addressing this issue have achieved heterogeneous results (Tunzelmann, Ranga, Martin, & Geuna, 2003). With reference to research performance of German BuSs, Dyckhoff et al. (2009) as well as Clermont, Dirksen, and Dyckhoff (2015) identify constant returns to scale for medium-sized departments, in fact. However, Clermont et al. (2015) also state that there are tendencies towards decreasing returns to scale for BuSs that consist of more than 16 to 19 professorships. Thus, this implicit CHE premise of constant returns to scale should be regarded critically.

CHE's diverse choices of cost categories in the form of staff size as a denominator for the relative performance ratios are unclear. CHE justifies neither its particular choices nor its criteria. But other cost categories would certainly be thinkable and plausible. For example, the relative PhD PPR (per professor) or the relative national publications PPR (per postdoctoral scientist) assume that either the contribution of the remaining research assistants is insignificant or that the ratio of professors/postdoctoral scientists to research assistants is constant among BuSs in general. However, it seems more plausible that both PhD and publication efficiencies also depend on the number of research assistants as co-producers. Regarding the assumption that the ratio of research assistants is constant, we can take a look at the professors and associated research assistants of each BuS. Accordingly, it can be shown that both the ratio for an individual BuS, intertemporally over time, as well as for different BuSs intratemporally within a particular ranking, vary considerably. Considering these relativizations, BuSs employing more research assistants ought to have competitive advantages over BuSs with fewer research assistants, at least in these relative PPRs. In contrast, third-party funds and international publications are set in relation to the total number of scientists, i.e. a certain contribution from research assistants is assumed here. But it is strongly questionable as to

what extent their contribution toward acquiring third-party funds and generating international publications would be different from their contribution to the writing of PhD dissertations and the generating of national publications.

As noted above, the basic benchmarking concept is quite transparent and easily understandable. According to CHE, when determining the factual top, middle, and lowest groups separately for each of the eight PPRs, significant differences in the data structure are taken into account by a moderate adjustment slightly downwards of the groups' transitions. Even though CHE argues that these actual frontiers are adjusted moderately upwards or downwards (Berghoff et al., 2011, p. B2), the overall context has the result in particular that those BuSs that are located slightly below the next group's transition are upgraded into the higher group (see also the explicit description in the CHE methods wiki (CHE, 2014)). But a detailed examination of the eight PPRs shows divergences from the described method in five of them. This, in turn, restricts the understandability and verifiability by external stakeholders as well as precluding a faithful representation, in terms of neutrality and freedom from error.

To illustrate this issue, the following two Tables 3 and 4 show extracts of the four absolute PPRs (Table 3) and the four relative PPRs (Table 4) from the CHE research ranking 2011. The first column in each table shows the ranking position, the second column serves as an index for each BuS, the third column presents the respective absolute or relative performance criterion's value, and the fourth column illustrates the criterion for the group classification, as described in section 4.1. For each performance criterion, the top group is highlighted in dark gray, the middle group in light gray, and the lowest group is without color.

**Table 3**  
*Extracts from the four absolute PPR results*

Position	BuS No.	National publications	Cumulative ratio	BuS No.	International publications	Cumulative ratio	BuS No.	PhDs	Cumulative ratio	BuS No.	Third-party funds	Cumulative ratio
1	46	178.8	4.70%	48	89.7	16.14%	52	52.3	5.62%	08	3912.7	5.45%
11	49	87.8	35.80%	51	11.7	51.34%	01	22.0	38.57%	02	2017.0	42.38%
12	35	84.0	38.01%	11	10.7	53.26%	33	21.3	40.86%	60	2001.0	45.17%
13	55	83.1	40.19%	69	10.0	55.06%	02	20.0	43.01%	01	1921.8	47.84%
14	30	82.0	42.35%	01	9.7	56.81%	35	20.0	45.16%	11	1880.3	50.46%
15	16	74.8	44.32%	06	9.3	58.48%	58	19.0	47.20%	16	1763.6	52.91%
16	37	71.4	46.20%	52	9.3	60.15%	12	18.3	49.17%	12	1665.2	55.23%
17	36	71.2	48.07%	35	9.0	61.77%	16	18.3	51.13%	57	1595.2	57.45%
18	06	70.8	49.93%	64	8.7	63.34%	34	17.7	53.04%	53	1568.8	59.63%
19	18	67.9	51.72%	16	8.0	64.78%	06	17.3	54.89%	38	1487.0	61.70%
20	24	64.5	53.41%	20	8.0	66.22%	55	17.3	56.75%	07	1456.1	63.73%
31	10	37.3	79.82%	27	4.0	88.77%	28	8.7	83.51%	14	560.0	89.76%
32	40	36.5	80.78%	29	4.0	89.49%	43	8.3	84.40%	25	511.9	90.47%
33	17	36.4	81.74%	38	4.0	90.21%	64	8.0	85.20%	28	479.2	91.14%
34	07	35.8	82.68%	47	4.0	90.93%	50	7.7	86.09%	59	478.7	91.81%
35	62	34.6	83.59%	28	3.7	91.66%	67	7.7	86.91%	54	441.5	92.42%
36	65	33.5	84.47%	59	3.7	92.26%	32	7.3	87.70%	67	397.8	92.98%
37	25	33.4	85.35%	39	3.3	92.86%	27	7.0	88.45%	62	393.2	93.52%
38	27	33.2	86.23%	68	3.3	93.45%	37	7.0	89.20%	18	381.9	94.05%
39	32	33.0	87.09%	63	3.0	93.99%	40	7.0	89.95%	64	310.6	94.49%
40	04	30.7	87.90%	32	2.7	94.48%	38	6.7	90.67%	27	305.7	94.91%
41	29	30.1	88.69%	36	2.7	94.96%	56	6.7	91.39%	37	301.5	95.33%
42	54	29.4	89.47%	58	2.7	95.45%	65	6.7	92.11%	50	299.4	95.75%
43	43	28.1	90.21%	62	2.7	95.93%	68	6.7	92.83%	09	292.0	96.15%
44	64	27.7	90.93%	05	2.3	96.35%	25	6.3	93.51%	45	288.8	96.56%
67	26	21.2	99.06%	41	0.3	99.95%	13	1.7	100.00%	32	98.3	100.00%
68	41	18.1	99.55%	56	0.3	100.00%						100.00%
69	47	17.8	100.00%									

Top group  
Middle group  
Lowest group

**Table 4**  
*Extracts from the four relative PPR results*

Position	BuS	National publications	Cumulative ratio (BuS)	BuS	International publications	Cumulative ratio (BuS)	BuS	PhDs	Cumulative ratio (BuS)	BuS	Third-party funds	Cumulative ratio (BuS)
1	48	4.7	1.45%	09	3.2	1.47%	08	3.8	1.49%	21	105.5	1.49%
14	42	3.2	20.29%	65	0.7	20.59%	62	1.5	20.90%	22	34.6	20.90%
15	24	3.1	21.74%	07	0.6	22.06%	21	1.4	22.39%	20	33.4	22.39%
16	27	3.0	23.19%	24	0.6	23.53%	35	1.4	23.88%	01	32.4	23.88%
17	35	3.0	24.64%	49	0.6	25.00%	46	1.4	25.37%	29	31.5	25.37%
18	40	3.0	26.09%	61	0.6	26.47%	48	1.4	26.87%	66	30.7	26.87%
19	46	3.0	27.54%	01	0.5	27.94%	63	1.4	28.36%	46	29.9	28.36%
20	16	2.9	28.99%	02	0.5	29.41%	16	1.3	29.85%	51	29.9	29.85%
50	55	2.0	72.46%	63	0.3	73.53%	55	0.7	74.63%	64	10.8	74.63%
51	02	1.9	73.91%	68	0.3	75.00%	24	0.6	76.12%	62	10.2	76.12%
52	64	1.9	75.36%	45	0.2	76.47%	25	0.6	77.61%	50	9.7	77.61%
53	17	1.8	76.81%	28	0.2	77.94%	38	0.6	79.10%	28	9.5	79.10%
54	23	1.8	78.26%	32	0.2	79.41%	41	0.6	80.60%	14	9.3	80.60%
55	59	1.8	79.71%	36	0.2	80.88%	42	0.6	82.09%	27	9.2	82.09%
56	67	1.8	81.16%	39	0.2	82.35%	54	0.6	83.58%	13	8.4	83.58%
57	09	1.7	82.61%	50	0.2	83.82%	67	0.6	85.07%	47	8.0	85.07%
58	61	1.7	84.06%	54	0.2	85.29%	05	0.5	86.57%	37	7.4	86.57%
67	32	1.3	97.10%	42	0.0	98.53%	22	0.1	100.00%	32	1.9	100.00%
68	43	1.3	98.55%	56	0.0	100.00%						
69	04	1.2	100.00%									

Top group  
Middle group  
Lowest group

An analysis of the extracts in Table 3 and 4 shows that the classification procedure based on the factual dataset of 2011 leads to diverging group sizes, especially with regard to the absolute PPRs. The more right-skewed the performance

criterion's distribution is, the fewer BuSs there are in the top and the middle group. Analogically, the more symmetric or left-skewed the performance criterion's distribution is, the more BuSs are allocated to these groups. Regarding the CHE method, it should be stated, however, that it leads to some limitations in terms of the interpretation of results and the explanatory power of the disclosed results, in particular to disadvantages for certain BuSs. In this respect, an oversensitivity of slight differences between values on group transitions should be avoided when classifying the groups. But as becomes obvious in Tables 3 and 4, values on transitions between two groups sometimes lie closely together. This can happen in spite of the existence of significant differences in the data structure, for example if there are no significant differences below the group transition (Rassenhövel, 2010, p. 183). As is shown in Table 3, the actual determination of group transitions appears to be quite arbitrary, since CHE does not consistently observe the principles it has itself established. According to these principles, BuS 43 with reference to national publications, BuSs 38, 56, 65, and 68 with reference to the number of PhDs, and BuS 16 with regard to the amount of third party funds should belong to the middle group. In addition, BuSs 28, 59, 54, 67, 62, 18, and 64 should be allocated to the lowest group. As a result, BuS 16 would then no longer be classified as being "Strong in research", as it is now. These discrepancies in the actual classifications are not verifiable, even if significant discontinuities in the data structure are taken into account. Let us illustrate this by looking at an actual numerical example using the third-party funds PPR: with reference to the top group of the actual PPR, the value of BuS 12 is 5.58% lower than that of BuS 16. If we apply the CHE principles strictly, BuS 16 would no longer belong to the top group, and the significance of the difference between the last value of the top group and the first value of the middle group would then be increased to 6.21% (BuS 16 versus BuS 11). The same results are obtained when regarding the transition from the middle and the lowest group (BuSs 64 to 27). At the moment there is a difference in percentage of 1.58%, which is much lower than the difference of 6.39% when CHE principles are applied strictly (then, between BuS 25 and BuS 28). Therefore,

understandability, consistency, and thus faithful representation, are quite limited.

Similar discrepancies are also observable within the relative PPRs (Table 4). Relative to the number of international publications, it is incomprehensible as to why BuSs 07, 24, 49, and 61 do not belong to the respective top group, particularly if we regard to homogeneous group classifications among the different PPRs. Moreover, some BuSs are allocated to different groups, although they exhibit the same value in a performance criterion. This is the case for BuSs 02 and 64 relative to national publications and for BuSs 45 and 28ff., relative to international publications.

With reference to the overall evaluation of research strength, Ahn et al. (2007) point out that the fundamental dominance principle known from decision theory (Keeney & Raiffa, 1993, p. 69f.; Eisenführ, Weber, & Langer, 2010, p. 95ff.) is neglected. Hence, considering the group of BuSs which are classified as being “Strong in research”, there are certain BuSs showing higher values in all eight evaluated performance criteria than other BuSs of the same group. In addition, the CHE benchmarking does not identify potential performance improvements and adequate benchmarking partners. Moreover, an equal weighting of all considered performance criteria is not necessarily in the interest of each external and university-internal stakeholder, since these may have different preferences regarding the importance of each cost and benefit category, and thus, certain criteria.<sup>11</sup> When awarding the designation “Strong in research”, the final equal weighting of the eight PPRs not least favors large BuSs, because they hold per se advantages with reference to the absolute PPRs, simply because of their size.

## SUMMARY AND IMPLICATIONS

The scope of our paper involved the development of a framework to analyze the CHE research ranking of German BuSs on the basis of fundamental accounting principles, and particularly by evaluating the extent of its compliance with fundamental IFRS accounting principles. We structured our analysis by using the basic concepts of cost and activity



accounting known from managerial accounting. Subsequently, we summarize the elementary structures of our framework and the resulting conclusions in section 5.1. In addition, we proffer concrete suggestions for improving this ranking system. In section 5.2, we illustrate that our analyses, which are specifically applied, are transferable to an evaluation of other rankings and we thus show the general usability of our suggested framework.

### ***Specific Implications and Recommendations for Improvement***

According to fundamental qualitative characteristics, the published information should be relevant with respect to the information requirements of the ranking recipients. In addition, the information should be represented faithfully, by ensuring the principles of completeness, neutrality, and freedom from errors. As discussed, the seven cost and benefit categories defined by CHE are definitely *relevant* in the context of evaluating and analyzing the performance of a BuS by generally fulfilling the functions of confirmation and prediction. The *completeness* of the considered cost and benefit category set can hardly be clarified conclusively. Whereas published company reports are meant to inform investors about the financial position of a firm, the information purpose of the CHE ranking is not distinct due to the heterogeneously addressed university stakeholders and their specific systems of objectives. Beyond that, CHE only evaluates those performances as cost and benefit categories which can be measured quantitatively. Further essential benefits are not or insufficiently covered (by proxy variables), e.g. the quality of young scientists' education. Indeed, not all possible cost and benefit categories discussed in the literature are considered by CHE, and this results in partial research performance evaluations. Due to the missing appropriate scientific studies, though, we can only speculate as to what extent additional cost and benefit categories might create new information about research performance and perhaps lead to divergent final evaluations. Indeed, there are first studies which reflect the relevance of diverse citation indicators (e.g. Waldkirch, Meyer, & Zaggel, 2013) as well as the impact of querying alternative literature databases on performance assessments of individual researchers in certain sub-disciplines of Business Administration

(Meyer, Waldkirch, & Zaggl, 2012). Detailed analyses of how far the implementation of alternative indicators in the CHE method might impact the performance evaluation are missing to date. In addition, it is not yet clarified how qualitative aspects of research performance might be considered, e.g. with regard to reviewed PhD dissertations. Furthermore, the usefulness of covering additional information should always – with reference to the cost constraint – exceed its costs. This, in turn, comes into conflict with the complexity of transferring certain cost and benefit categories, mentioned above, into measurable indicators and therefore possible extensions of the category set.

The two principles *neutrality* and *correctness* are directly interconnected to the postulated qualitative characteristic of *verifiability* of published information. In this regard, and due to the clearly existing principal agent situation, especially the data reported by a BuS has to be scrutinized critically. While the principal (either CHE or its information recipients) is interested in faithfully represented, intertemporally and intratemporally comparable, useful information for decision making, the agent (the BuS) has the extrinsic motivation to be presented as favorably as possible in comparison with its peer group. Potential divergent interests and asymmetrically distributed information concerning the factual research performance of a BuS in combination with the de facto impossibility of precisely verifying the reported data might result in personnel coordination problems. However, also data acquired by CHE itself – national and international publications – are not verifiable for external recipients, because CHE does not provide raw data.

Another qualitative characteristic addresses the question of how far the published information is intertemporally and intratemporally *comparable*. Abstracting from misguided incentives and assuming consistent objectives, the intratemporal comparability, meaning the possibility to compare the performance of different BuSs within the same period, might still be limited. For example, this might be caused by the lack of clearly worded or defined questions in the CHE questionnaire, resulting in false reporting or reporting based on varying assumptions by the BuSs. This, of course, limits the comparability of information. The intertemporal comparability

of individual BuSs in the course of consecutive time periods is limited, too, because CHE has altered its data acquisition method as well as its definition of underlying cost and benefit categories over the years. CHE, however, implemented these modifications in order to enhance its ranking method and should naturally not be judged negatively for having done so. However, it is imperative that information recipients consider these changes when analyzing the data in intertemporal studies.

With regard to the principle of *timeliness*, we have to question the underlying reporting period of three years. The shorter a reporting period, the more volatile the ranking results will certainly be. Longer reporting periods will reduce random fluctuations of performance but timeliness and accordingly relevance of information will simultaneously decrease. Thus, the vital point consists of determining a time period which ensures a balanced relation between timeliness and reliability. Basically, three years would appear to be reasonable. A one-year time period as the decisive period in the context of financial reports would seem to be too short with regard to the unpredictability of benefits in university research, e.g. the time period between submitting and having a scientific journal article published is hard to predict due to the prior peer-review process. Finally, determining the most appropriate time period more or less depends on each individual cost and benefit category. For example, three years would tend to be too short for the evaluation of citations as an impact indicator (Research Evaluation and Policy Project, 2005, p. 20f.).

Concerning the qualitative characteristic of *understandability*, we can state that the design and procedure of the CHE ranking are per se easily and intuitively comprehensible. Thus, the concise presentation can, in principle, provide informative insights into the research profile of a BuS and its performance. The disclosure of individual values in the PPRs enables the ranking audience further opportunities for their own specific analyses and evaluations. Nevertheless, adequately interpreting the disclosed information postulates that the recipient has substantial knowledge in terms of discipline and research environment. Otherwise, misleading implications might be deduced. As illustrated above, possible misinterpretations also

result to some extent from incorrect, imprecise, and contradictory descriptions of the detailed ranking method.

In conclusion, it should be stated that many problems and misguided incentives resulting from the CHE rankings seem to be promoted by the fact that the rankings represent – in terms of financial accounting – flow measures rather than stock measures. In this respect, six of the seven cost and benefit categories are periodic key figures; only the number of postdoctoral scientists constitutes a cumulative key figure on a specific date. This implies that the usage of diverse and inconsistent cost allocation principles as well as taking advantages of discretionary freedom are not balancing over consecutive periods and, accordingly, the total time period. In fact, a BuS can even allocate its costs and benefits differently from past reporting with respect to an advantageous positioning in the various academic disciplines, e.g. by multiple listings of (research-strong) interdisciplinary professorships.

Due to our previous analyses, we are now able to derive several recommendations for improvement which CHE could easily implement in order to mitigate some of the problems and unintended incentives. By applying these, acceptance and significance of the research ranking could be promoted and any lost confidence could be regained.

- Limiting the discretionary freedom of a BuS: The previously described discretionary freedom in the data that a BuS provides should be reduced as far as possible. Specifically, formulations in the questionnaire should be expressed or defined clearly so that there is no leeway for interpretation. For example, CHE should define distinctly which interdisciplinary professorships are to be allocated to a BuS and which not, and at what precise time stage a PhD degree has actually been achieved from a reporting perspective.
- Disclosure of raw data: CHE should disclose the underlying raw data of its evaluation in order to create transparency and to enable information recipients to clarify potential discrepancies, e.g. the postdoctoral scientists' names that are surveyed in the context of the bibliometric analyses. On the one hand, this could also lead to corrections of accidentally

false data by critical users, on the other hand it might prevent possible incentives to miss-report.

- Detailed explanation of the ranking method: Currently, CHE does not document all information regarding data and procedure. But this can be obtained in part by consulting further CHE publications. Accordingly, recipients also have to examine, for example, CHE working papers, the methods wiki on the CHE website, or published questionnaires. Otherwise, they have to clarify ambiguities by contacting CHE directly. However, the information is incomplete and there still exist different and conflicting descriptions of procedures. CHE should consolidate all essential updated information about the method and verify the compliance to the qualitative principles mentioned in section 2, especially to the principles of completeness, freedom from errors, and timeliness.
- Interactive interpretation option of cost and benefit categories: In order to address different CHE stakeholders' objectives, the option of executing individual interactive calculations and interpretations could be implemented, e.g. on the CHE website. Similar options already exist for the CHE university ranking that focuses on academic teaching.

### ***General Implications***

This paper was primarily focused on the specific evaluation of the CHE research ranking of German BuSs. The transferability of our findings, however, is reflected by considerable parallels between different ranking procedures and their respective critique. For instance, with reference to the college ranking of the US News and World Report, the completeness and relevance of the considered indicators for evaluating education performance is discussed, too (e.g. Carey, 2006). In analogy to CHE, several of these indicators are acquired by surveying the colleges, which is why problems of verifiability and correctness of data also apply here. In this context, "since the mid-1990s numerous stories in the popular press have documented how various schools distort their standard operating procedures, creatively interpret survey instructions, or boldly misreport information", e.g. their SAT

scores, “in order to raise their rankings” (Diver, 2005). These problems are basically valid for the doctoral program ranking of the National Research Council as well, since it also incorporates information reported by the universities themselves. Moreover, university publications are covered by querying the WoS database, so that a majority of the problems involved in the CHE research ranking are valid here, too.<sup>12</sup>

Hence, it becomes obvious that our structure of analyzing university rankings on the basis of accounting principles is also applicable to other rankings. Nearly every university ranking aims to create transparency concerning university performance by supplying information to university stakeholders. Due to the fact that the main function of accounting is the supply of useful information on companies for decision making, particular accounting principles are certainly predestinated to be a kind of theoretical basis for evaluating and designing rankings. For decades, accounting has been engaged in developing adequate principles and procedures in order to generate, process, and publish financial information on companies. These principles have been developed and advanced over time and have proven to be reasonable for informing both internal and external company stakeholders (Alexander & Nobes, 2013, p. 65f.). In addition, diverse incentives originating from accounting are extensively analyzed in the literature by theoretical as well as empirical studies (e.g. Gao, 2013; Cai, Rhaman, & Courtenay, 2014), leading to incorporations of relevant insights into developments or advancements of accounting principles. As shown in this paper, the presented IFRS principles of the *lex generalis* are appropriate for defining basic ranking requirements; regulations for the detailed ranking design as some kind of *lex specialis*, however, need additional research. The existing *lex specialis* of IFRS is certainly not applicable for usage in the context of university rankings and would have to be developed specifically.

Ultimately, such enhancements could possibly lead to universities participating in precise and systematic provision or disclosure of information to interested stakeholders. Against this background it is basically incomprehensible as to why certain information, e.g. concerning spending of public funds, is not or

cannot be disclosed to stakeholders by universities and their departments. Therefore, some authors suggest the implementation of an “academic accounting system”, where universities or their departments systematically record, calculate, and publish their academic/scientific costs and benefits (e.g. Kirchhoff-Kestel, 2006; Küpper, 2013). A corresponding result could then consist of a comprehensive and regularly conducted university and department reporting that satisfies both the information needs of the stakeholders and the reduction of the departments’ administrative efforts to a reasonable degree. However, the complexity and multidimensionality of a university’s mission complicates such a simple evaluation of academic/scientific costs and especially benefits. This might be the reason why the German Council of Science and Humanities<sup>13</sup> – in cooperation with professional representatives, especially the relevant associations – is working on a so called “core dataset research”. The purpose is to acquire research data discipline-specific and nationwide in order to enable interested stakeholders, e.g. ranking institutions, to access this raw data (German Council of Science and Humanities, 2013). On the basis of such a generally accepted dataset, further, substantially empirical analyses of research activities could be facilitated. With regard to these efforts, the suggested accounting principles could be helpful to derive basic requirements for such information systems.

## REFERENCES

- Ahn, H., Dyckhoff, H., & Gilles, R. (2007). Data aggregation for performance assessment using ranking: Comparison of the CHE and DEA methodology as well as development of a compromise solution. *Zeitschrift für Betriebswirtschaft*, 77(6), 615–643 [in German].
- Alexander, D., & Nobes, C. (2013). *Financial accounting: An international introduction*. 5<sup>th</sup> edition. Harlow, England: Pearson.
- Backes-Gellner, U., & Zanders, E. (1989). Lehre und Forschung als Verbundproduktion: Data-Envelopment-Analysen und organisationsökonomische Interpretationen der Realität in wirtschaftswissenschaftlichen Fachbereichen. *Zeitschrift für Betriebswirtschaft*, 59(3), 271–290.
- Baum, J. A. C. (2011). Free-riding on power laws: Questioning the validity of the impact factor as a measure of research quality in organization studies. *Organization*, 18(4), 449–466.
- 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, Germany: Centrum für Hochschulentwicklung.
- Boor, M. (1982). The citation impact factor: Another dubious index of journal quality. *American Psychologist*, 37(8), 975–977.
- Bort, S., & Schiller-Merkens, S. (2010). Publish or perish: Was die Bedeutung organisatorischer Konzepte mit dem Leistungsdruck im Wissenschaftsbetrieb zu tun hat. *Zeitschrift Führung und Organisation*, 79(5), 340–346.



- Cai, L., Rahman, A., & Courtenay, S. (2014). The effect of IFRS adoption conditional upon the level of pre-adoption divergence. *The International Journal of Accounting*, 49(2), 147–178.
- Carey, K. (2006). *College rankings reformed: The case for a new order in higher education*. Washington D.C., USA: Education Sector Reports.
- Chalmers, A. F. (1990). *Science and its fabrication*. Buckingham, England: Open University Press.
- CHE (2014). MethodenWiki. <http://www.che-ranking.de/methodenwiki/index.php/Hauptseite> (2014-02-20).
- Clermont, M., & Schmitz, C. (2008). Acquisition of economic journals in the ISI-databases and in Scopus. *Zeitschrift für Betriebswirtschaft*, 78(10), 987–1010 [in German].
- Clermont, M., & Dyckhoff, H. (2012). Acquisition of economic journals in literature databases. *Betriebswirtschaftliche Forschung und Praxis*, 64(3), 324–346 [in German].
- Clermont, M., Dirksen, A., & Dyckhoff, H. (2015). Returns to scale of Business Administration research performance in Germany. *Scientometrics*, 103(2) 583-614.
- Diver, C. (2005). Is there life after rankings? *The Atlantic Nov. 1<sup>st</sup>*, 2005. <http://www.theatlantic.com/magazine/archive/2005/11/is-there-life-after-rankings/304308/2/> (2014-07-25).
- Dyckhoff, H. (2006). *Produktionstheorie: Grundzüge industrieller Produktionswirtschaft*. 5<sup>th</sup> edition. Berlin, Germany: 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., Gilles, R., & Schmitz, C. (2005). Beurteilung der Forschungsleistung und des CHE-Forschungsrankings betriebswirtschaftlicher Fachbereiche. *Wirtschaftswissenschaftliches Studium*, 34(2), 62–69.

- 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.
- Eisend, M. (2011). Is VHB-JOURQUAL2 a good measure of scientific quality? *Business Research*, 4(2), 241–274.
- Eisenführ, F., Weber, M., & Langer, T. (2010). *Rational Decision Making*. Heidelberg, Germany: Springer.
- Ewert, R., & Wagenhofer, A. (2014). *Interne Unternehmensrechnung*. 8<sup>th</sup> edition. Berlin, Germany: Springer.
- 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.
- Frey, B. S. (2007). Evaluierungen, Evaluierungen ... Evaluitis. *Perspektiven der Wirtschaftspolitik*, 8(3), 207–220.
- Gao, P. (2013). A measurement approach to conservatism and earnings management. *Journal of Accounting and Economics*, 55(2), 251–268.
- German Council of Science and Humanities (2013). *Empfehlungen zu einem Kerndatensatz Forschung*. Berlin, Germany: German Council of Science and Humanities.
- Hornbostel, S. (2001). Third party funding of German universities: An indicator of research activity? *Scientometrics*, 50(3), 523–537.
- International Accounting Standards Board (2010). *Conceptual framework for financial reporting 2010*. London, England: IFRS Foundation Publications Department.
- Jansen, D., Wald, A., Franke, K., Schmoch, U., & Schubert, T. (2007). Drittmittel als Performanzindikator der wissenschaftlichen Forschung: Zum Einfluss von Rahmenbedingungen auf Forschungsleistung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 59(1), 125–149.

- 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 metrics? *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.
- Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value tradeoffs*. Cambridge, England: Press Syndicate of the Cambridge University.
- Kieser, A. (2012). JOURQUAL: Its use, not its misuse, is the problem. Or: Why Wirtschaftsinformatik is the best journal of the German-speaking countries. *Die Betriebswirtschaft*, 72(2), 93–110 [in German].
- Kirchhoff-Kestel, S. (2006). *Kosten- und Leistungsmanagement in Hochschulen: Grundlagen und Konzepte für ein zweckorientiertes Rechnungswesen*. Lohmar, Germany: Eul.
- Küpper, H.-U. (2013). A specific accounting approach for public universities. *Journal of Business Economics*, 83(7), 805–829.
- 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.
- Meyer, M., Waldkirch, R. W., & Zaggel, M. A. (2012). Relative performance measurement of researchers: The impact of data source selection. *Schmalenbach Business Review*, 64(4), 308–330.
- Rassenhövel, S., & Dyckhoff, H. (2006). Die Relevanz von Drittmittelindikatoren bei der Beurteilung der Forschungsleistung im Hochschulbereich. In: S. Zelewski & N. Akca (Eds.), *Fortschritt in den Wirtschaftswissenschaften: Wissenschaftstheoretische Grundlagen und exemplarische Anwendungen* (pp. 85–112). Wiesbaden, Germany: Gabler.

- Rassenhövel, S. (2010). *Performancemessung im Hochschulbereich: Theoretische Grundlagen und empirische Befunde*. Wiesbaden, Germany: Gabler.
- Reedijk, J., & Moed, H. F. (2008). Is the impact of journal impact factors decreasing? *Journal of Documentation*, 64(2), 183–192.
- Research Evaluation and Policy Project (2005). *Qualitative indicators for research assessment: A literature overview*. Canberra, Australia: The Australian National University.
- Ruhnke, K., & Simons, D. (2012). *Rechnungslegung nach IFRS und HGB*. 3<sup>rd</sup> edition. Stuttgart, Germany: Schäffer Poeschel.
- Schrader, U., & Hennig-Thurau, T. (2009). VHB-Jourqual2: Methods, results, and implications of the German 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, Belgium: European University Association.
- Tunzelmann, N. von, Ranga, M., Martin, B., & Geuna, A. (2003). *The effects of size on research performance: A SPRU review*. Brighton, England: University of Sussex.
- Usher, A., & Savino, M. (2006). *A world of difference: A global survey of university league tables*. Toronto, Canada: Educational Policy Institute.
- Waldkirch, R. W., Meyer, M., & Zaggl, M. A. (2013). Beyond publication counts: The impact of citations and combined metrics on the performance measurement of German business researchers. *Zeitschrift für Betriebswirtschaft, Special Issue 3/2013*, 61–86.
- Weber, M., Henderson, A. M., & Parsons, T. (1964). *The theory of social and economic organization*. London, England: Collier-Macmillan.

## NOTES

1. Since, as far we know, CHE does not intend to implement any modifications to the method for the next BuS research ranking in 2014, our analyses will also be valid for prospective rankings.
2. In this context, however, it should be noted that CHE does use the adjective “visible” in one text passage in the ranking publication available on the CHE website (Berghoff et al., 2011, p. E15). CHE is apparently aware of this misleading terminology but does not use the more precise wording consistently.
3. We refer to the issue of this classification below in section 4.
4. Beasley (1990), Beasley (1995), and Fandel (2007), for instance, use third-party funds in their empirical studies both as input and output. A detailed analysis concerning this issue is provided by Rassenhövel and Dyckhoff (2006).
5. A “Habilitation” is a specific post-doctoral dissertation that qualifies a person for a professorship in Germany.
6. In this context, we tried to gather information on titles and authors of published PhD dissertations by requesting the BuSs directly to verify the numbers disclosed in the CHE ranking. When doing so, we encountered resistance from some BuSs who were either not willing or unable to provide this information.
7. We do not mention the explicit BuS names on purpose, because we do not want to criticize any BuS. Our intention rather is to identify and illustrate possible inconsistencies in the CHE research ranking.
8. Meanwhile, CHE itself takes this circumstance into account by emphasizing the multi-dimensionality of academic performance through the publication of a new, enhanced ranking called “Manifold Excellence”. Apart from research performance, three other dimensions are also highlighted in this ranking: practical orientation of teaching, internationality of teaching and research, and student orientation.
9. While our previously discussed statements, e.g. on error-free and unbiased acquisition and publication of data, are valid for all university stakeholders, the question of the comparability of BuSs certainly depends on the specific objectives of the different ranking users. The considerations mentioned in this paragraph are mainly important from the perspective of the ranked BuSs especially if they are assessed by university management on the basis of their ranking results. For instance, from a prospective student’s point of view, the requirements shift due to varying objective function weights. Hence, this group is more interested in the net productivity of the total service demands of a

BuS rather than in the absolute productivity of the research performance of that BuS.

10. Backes-Gellner and Zanders (1989, p. 227ff.), for example, discovered that BuSs that belong to established traditional universities have efficiency advantages with regard to their teaching and research production.

11. Indeed, it is hardly possible to establish a generic ranking scheme which takes all different stakeholder interests into consideration without being too complex.

12. It should be positively noted that, in contrast with CHE, the National Research Council requests publication lists from the ranked universities, enabling verification of the articles by querying the WoS.

13. "The German Council of Science and Humanities (..) provides advice to the German Federal Government and the State (..) Governments on the structure and development of higher education and research" (see <http://www.wissenschaftsrat.de>).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.