

ACCOUNTING RESEARCH PRODUCTIVITY BENCHMARKS: WHO USES THEM?

Wm. Dennis Huber, JD, LL.M., DBA
Capella University
225 South 6th Street
Minneapolis, MN 55402
Dennis.Huber@Capella.edu

James A. DiGabriele, Ph.D.
Montclair State University
1 Normal Ave
Upper Montclair, NJ 07043
digabrielej@mail.montclair.edu

Keywords: Accounting Research Productivity, Benchmarks

ACCOUNTING RESEARCH PRODUCTIVITY BENCHMARKS: WHO USES THEM?

ABSTRACT

Numerous studies have been conducted over the last 40 years concerning accounting faculty research productivity that have developed benchmarks for which accounting department administrators can evaluate the research productivity of accounting faculty. Such benchmarks not only use quantity of publications, but also subjective qualitative measures such as the perceived “quality” of journals in which the research is published. It remains unknown, however, the extent to which, or whether, any of the benchmarks published to date are actually used by administrators to evaluate the accounting research productivity of their faculty or what decisions they may make based on those benchmarks. The research question is thus, do accounting department administrators use published accounting productivity benchmarks.

This study surveyed accounting department administrators (deans, chairpersons, and heads of accounting departments) of U.S. institutions to determine if they use published accounting research productivity benchmarks in evaluating faculty for hiring, promotion, tenure, or merit pay decisions. The results of our survey were analyzed using chi-square tests of association and one-sample chi-square. The results indicate that at a .000 level of significance accounting department administrators do not use published accounting research productivity benchmarks in evaluating faculty for hiring, promotion, tenure, and merit pay decisions. This suggests that new research productivity benchmarks need to be developed that can actually be used by administrators.

INTRODUCTION

More than 30 studies have been conducted over the last 40 years concerning accounting research productivity that have purported to develop benchmarks¹ by which accounting department administrators can evaluate the research productivity of accounting faculty. More recently published benchmarks typically use not only the quantity of research articles published, but also quality measures such as the perceived rankings of journals in which the research is published (e.g., Hasselback, Reinstein and Schwan 2000; Richardson and Williams 1990).

¹ Research in accounting faculty productivity is not consistent in its use of terminology. Even though it is not accurate, “Benchmarks” is the most commonly used term, but other terms such as “metrics,” “productivity,” and “output” have also been used. Here, the term “benchmarks” is used to refer to all accounting research productivity measures.

Although most accounting faculty research productivity studies focus on the research productivity of accounting faculty of American universities, measuring accounting faculty research productivity is by no means limited to American academia. Richardson and Williams (1990), for example, studied Canadian accounting faculty research productivity. Their study parallels those of American faculty research productivity studies in identifying the research productivity of Canadian accounting faculty in Canadian universities. Their findings of journal rankings are consistent with studies of American accounting faculty research productivity studies using journal rankings. Similarly, Chan, Chen and Cheng (2005) studied the research productivity of accounting for Asia-Pacific universities. Their findings also found the research productivity to be comparable to leading universities in North America.

According to the researchers, their benchmarks can be used by administrators in hiring, promotion, tenure, and merit pay decisions. It remains unknown, however, the extent to which, or whether, any of the benchmarks published to date are actually used by administrators to evaluate the accounting research productivity of their faculty or what decisions they may make using published accounting research productivity benchmarks.

This study surveyed accounting department administrators to determine if they use published accounting research productivity benchmarks in evaluating faculty for hiring, promotion, tenure, and merit pay decisions. The results of the survey indicate that overall administrators do not use published accounting research productivity benchmarks in evaluating faculty for hiring, promotion, tenure, and merit pay decisions thus calling into question the value of published accounting research productivity benchmarks.

IMPORTANCE AND CONTRIBUTION

Evaluating accounting faculty performance is important for both faculty and administrators (Fogarty 2004). Faculty research productivity is believed to heavily influence hiring, promotion, tenure, and merit pay decisions (Hasselback, Reinstein and Abdolmohammadi 2012). Doctoral granting institutions normally place greater emphasis on publication activity than non-doctoral

granting institutions in evaluating faculty performance (Reinstein and Calderon 2006), and AACSB accredited institutions require publications for promotion (Kerr, Simkin and Mason 2009).

The problem presented for all institutions that use faculty research productivity as a factor in evaluating faculty performance, however, is how to evaluate faculty research productivity. Objective evaluation metrics are necessary in order to prevent biased decisions and cronyism (Walker, Fleishman and Stephenson 2013), and in order to provide to faculty the information they need to understand how their research productivity will be evaluated. (Hasselback and Reinstein 1995a; Hasselback, Reinstein and Schwan 2000; Walker, Fleishman and Stephenson 2013). Yet, every benchmark other than simple quantity of publications introduces subjective factors into productivity measures such as perceptions of institutional or program ranking or the quality of the journal in which the research is published (Hasselback, Reinstein and Schwan 2000; Richardson and Williams 1990).

Many benchmarks have been published over the last four decades that attempt to provide some measure of evaluating faculty research productivity. Yet, it is unknown whether administrators actually use published research productivity benchmarks in evaluating faculty research productivity as suggested by the benchmark studies.

This paper attempts to determine whether administrators actually do use published research productivity benchmarks to evaluate faculty research productivity in hiring, promotion, tenure, and merit pay decisions as suggested by the benchmark studies.² If administrators use published research productivity benchmarks for hiring, promotion, tenure, and merit pay decisions (Walker, Fleishman and Stephenson 2013), then the purposes for which they use published research productivity benchmarks and how they are used should be understood by faculty.

² Hiring, promotion, tenure, and merit pay decisions may be determined by committees. However, the published benchmark studies suggest that their benchmarks can be used by administrators rather than by committees for such decisions. This study therefore does not consider decisions by committees but only by administrators in keeping with the suggestions of the benchmark studies.

On the other hand, the failure of administrators to use published research productivity benchmarks to evaluate faculty research productivity calls into question whether the published research productivity benchmarks are useful for evaluating faculty research productivity. If they are not used, then it is incumbent on researchers and administrators to develop benchmarks that can be used in evaluating faculty research productivity for hiring, promotion, tenure, merit pay, and resource allocation decisions.

More importantly, however, is that this study reveals that the term “benchmark” in reference to measurements of faculty research productivity has for decades been misused and misunderstood. A benchmark is “a point of reference from which measurements may be made,” or, “something that serves as a standard by which others may be measured or judged.” (Merriam-Webster 2016).

Benchmarks based on relative performance can legitimately be used when those benchmarks are based on objective quantitative data to establish a point of reference or standard. One such example is an interest rate benchmark. “Yields on bonds issued by the U.S. treasury are viewed as providing a benchmark interest rate relative to which all other borrowers pay a default risk premium.” (Khurana and Raman 2003). Unlike accounting research productivity benchmarks, benchmarks such as interest rates are not subject to changes based on perceptions or opinions but on observable data.

However, the fact that accounting research productivity benchmarks are based on perceptions of journal quality or program rankings removes any pretense of objectivity by introducing subjective perceptions into the benchmarks. This subjectivity is recognized by every published benchmark that is based on perceptions of journal quality or program rankings. The problem of ranking as a function of subjectivity is compounded by the problem that rankings for accounting education authors and departments from Australia, Canada, New Zealand, the Republic of Ireland, the United Kingdom, and the United States yield significantly different results depending on the methodologies used to establish the rankings. (Bernardi, Zamojcin, and Delande 2016).

Thus, accounting research productivity benchmarks are not true benchmarks since they not only are not based on observable data, standards, baselines, or even methodology, but are constantly moving based on changes in perceptions. Nevertheless, the term benchmark is used here in order to be consistent with the prior studies.

No published faculty research productivity benchmark of the last 20 years establishes an objective standard by which faculty productivity may be measured or judged.³ With no objective standard or fixed point of reference faculty productivity benchmarks cannot be justifiably used to measure faculty research productivity since a faculty member would not know when or if the standard has been met, especially since the proposed benchmark is subject to change at any given moment as a result of changes in perceptions of journal quality or program rankings. However, the term benchmark is used in this study in order to be consistent with its usage in accounting research productivity studies.

REVIEW OF BENCHMARKS

More than 30 studies have been published over the past forty years that have developed various benchmarks that the researchers recommend, propose, or suggest can be used by accounting department administrators (chairs, deans, etc.) to evaluate the research productivity of accounting faculty. Many benchmarks incorporate rankings of journals, accounting doctoral programs, or universities. Some focused on the research productivity of individually named researchers in order for researchers to compare their own research productivity with the research productivity of other researchers (e.g., Hasselback, Reinstein and Schwan. 2003).

Each successive study of accounting research productivity purports to be an improvement over previous published benchmarks by adding a new dimension such as adding quality measures to quantity measures; for example, adding a time factor, or expanding a previously published benchmark from a limited number of schools or journals to include more schools or journals.

³ The University of Wyoming benchmarks developed benchmarks based on a point system. The UWy benchmarks are discussed below.

Many acknowledge that developing productivity metrics and benchmarks is difficult. Almost all caution against the potential abuse of productivity metrics and benchmarks for evaluating faculty research productivity, but nevertheless then go on to recommend that administrators and faculty can use their benchmarks for evaluating accounting faculty research productivity for hiring, tenure, promotion, and merit pay decisions thus ignoring their own warnings.

For this study, twelve benchmark studies published over the past 20 years were chosen for four reasons. First, it is extremely unlikely that administrators would be familiar with the more than 30 accounting research productivity benchmarks published over the last 40 years, and more unlikely that they would be familiar with older studies. Second, earlier studies were rather rudimentary, consisting of little more than merely counting the number of publications of accounting faculty published in various journals. Third, several benchmarks were published prior to the emergence of what are considered to be “elite” accounting journals.⁴ Publishing in the “elite” journals is seen as important by both faculty and administrators and publishing in the “elite” journals is included as a component of more recently published research productivity benchmarks. (Walker, Fleischman and Stephenson 2010).

Fourth, the primary purpose of several studies was more focused on establishing the perceived ranking of journals, programs, or institutions, or on the historical research productivity of individually named faculty members, rather than on establishing benchmarks that can be useful to administrators in evaluating faculty research productivity. Perceived ranking of journals or institutions, or the historical research productivity of individually named faculty members, without more, provides insufficient guidance for how to apply such data for forming actual benchmarks⁵ and are therefore excluded.

⁴ *The Accounting Review, Journal of Accounting Research, Journal of Accounting and Economics, Contemporary Accounting Research, and Accounting, Organizations and Society* are considered to be “elite” accounting research journals. (See, e.g., Richardson and Williams 1990)

⁵ For a more comprehensive review of published accounting research productivity benchmarks see Huber 2016a and 2016b.

Studies of accounting research productivity invariably recommend how their productivity benchmarks can be used by administrators, but their recommendations merely mirror the recommendations of previous studies. That is, it is claimed that the benchmarks of a particular study can be used for evaluating faculty research productivity for hiring, tenure, promotion, and merit pay decisions. Subsequent studies then echo the claim that their benchmarks can be useful for evaluating faculty research productivity for hiring, tenure, promotion, and merit pay decisions. There is thus an implicit suggestion that the benchmarks published later have greater usefulness for the same decisions than benchmarks established earlier, particularly since studies published later are motivated by what are seen as omissions or shortcomings in the earlier studies, which in turn implies that the earlier published benchmarks should be abandoned in favor of the later published benchmarks.

More important, however, is the fact that the published benchmarks are inconsistent. Since many published benchmarks incorporate surveys of perceived rankings of journals, programs, or universities as proxies for quality of research, the benchmarks change with changes in perceptions of the perceived rankings of journals, programs, or universities.

SUMMARY OF PUBLISHED BENCHMARKS

The purpose of this paper is not to critically analyze or evaluate the various published benchmarks. Nor is it to identify problems associated with using any particular set of published benchmarks for evaluating faculty research productivity. A formal literature review and critical analysis of each of the published benchmarks of faculty research productivity would therefore distract from the limited purpose of this study which is to determine whether administrators actually use published accounting research productivity benchmarks to evaluate faculty research productivity and if so, whether administrators use the benchmarks for the decisions recommended by the researchers. Instead, a concise summary of each published benchmark included in this study beginning with its identification of the weaknesses or shortcomings of previous published benchmarks, the basis and method used to establish the benchmarks,

highlights of what the benchmarks consist of, and their use as recommended by the authors of the benchmarks is presented.

The twelve benchmarks summarized here are limited to those that have been published for accounting faculty in American universities. There are several reasons for limiting the benchmarks for accounting faculty in American universities. First, more benchmarks have been published to measure research productivity of accounting faculty in American universities than to measure research productivity of accounting faculty in universities in other countries or regions. Second, the number of accounting faculty is much greater in American universities than in other countries or regions. Third, the number of universities is much greater in the United States than in other countries or regions. Fourth, the benchmarks published for accounting faculty in American universities may not be relevant to accounting faculty in universities in other countries or regions. The summary of the benchmarks is abridged from the respective publication of the benchmark.

Hasselback and Reinstein (1995a)

Identification of weaknesses of previous published benchmarks. Previous studies used only three methods to assess faculty research productivity: citation analysis, faculty/administration surveys, and counting the number of articles. No previous study ranked accounting doctoral programs using both qualitative and quantitative factors or considered the research records of all of their graduates over an extended period of time.

Basis of benchmarks. Benchmarks are based on the quantity and perceived quality of publications of 2,708 accounting graduates at 73 doctoral granting institutions based on 41 journals.

Method. The ranking of accounting doctoral programs is based on quantity and perceived quality and of publications of graduates of doctoral granting institutions

Summary of benchmarks. Graduates of larger public institutions dominate the total number of articles published. Graduates of private institutions dominate the total number of articles published when the results are weighted to consider the total number of graduates produced.

Suggested use. Potential faculty can use the results to gain a better understanding of the effectiveness of particular research programs. Administrators can use these results to make more informed decisions of faculty hires and to help establish reasonable standards for promotion and tenure decisions. Administrators can use the results to compare their graduates' productivity to that of peer institutions. Administrators can use the results to more efficiently recruit faculty, allocate resources, and direct program emphases.

Hasselback and Reinstein (1995b)

Identification of weaknesses of previous published benchmarks. Researchers have used only three methods to assess faculty research productivity: citation analysis, faculty/administration surveys, and counting the number of articles. Previous studies analyzed the quantity of articles that published by accounting faculty or the quality of journals where accounting faculty publish. None have combined both methodologies.

Basis of benchmarks. Assesses both the quality and quantity of the publication records of accounting faculty members at 716 institutions using 40 journals.

Method. Counted quantity of publications of accounting faculty and identifies individual accounting faculty by name and quantity of publications. Assigned weights to journals to determine quality.

Summary of benchmarks. Larger institutions granting accounting doctoral degrees dominated the highest rankings when looking at total number of articles. High quality private institutions with small numbers of faculty garnered the highest rankings. This result remained consistent when considering only the five elite accounting journals.

Suggested use. Faculty members, students, administrators, and alumni can use these results as a measure of the research records of these institutions.

Hasselback, Reinstein and Schwan (2000)

Identification of weaknesses of previous published benchmarks. Previous researchers have used three techniques to assess faculty research productivity: counting, citation analysis, and surveys. Counting does not provide measures of the quality of faculty research. When using counting, problem is whether to give publication credit to the faculty member's present institution or to the institution when the article was written.

Basis of benchmarks. Analyzed the research productivity of 3,289 faculty who graduated from accounting doctoral programs between 1971 and 1993.

Method. Ranked over 100 journals. Selected the highest rated journals.

Summary of benchmarks. Very few faculty have published in the Best 4 premier journals. One published article in a top four journal is likely to put its author in the top 20%, or even top 10%, of all faculty. After holding a doctorate for 10 years, it takes roughly four to six articles in 12 best journals to reach the top 10%, but only about two to reach the top 20%. For the critical periods of tenure and promotion an institution desiring to place in the top third of all institutions could set a benchmark of two or three articles in the best 40 journals.

Suggested use. These benchmarks are likely to be useful primarily for the highest rated institutions, such as Chicago or Stanford, which are likely to expect the publication records of their faculty to be in the top 10%. Administrators can use up to four parameters to select appropriate benchmarks for specific individuals, including whether to give the faculty member full credit for co-authored articles; the appropriate level of journal quality; the appropriate level of performance; and the number of years that a faculty member has been "in grade." This initial

set of benchmarks provides administrators with justification for specifying a required number of articles for tenure and promotion.

Christensen, Finger and Latham (2002)

Identification of weaknesses of previous published benchmarks. Previous accounting faculty productivity studies commonly omit publications in nonaccounting journals. Most previous accounting faculty productivity studies merely counted articles over a longer period of time.

Basis of benchmarks. Based on two groups of new scholars who received an accounting Ph.D. (316 in 1987/88 group and 253 1977/78 group) and took tenure-track positions at U.S. universities. Publications are based on 33 accounting journals; 141 non-accounting business journals.

Method. Journals were classified as either “practitioner” or “academic.”

Summary of benchmarks. The average publication count per 87/88 (77/78) scholar was 1.95 (1.55). The 87/88 scholars employed at Ph.D.-granting schools published on average, 2.72 articles per scholar in accounting journals compared to 1.32 articles per scholar for those employed at non-Ph.D.-granting schools. For the 87/88 sample, there was an average productivity of 1.95 total publications, but only 1.38 publications in accounting journals only. The 77/78 average for Ph.D. vs. non-Ph.D.-granting institutions was 2.54 vs. 0.53 articles per scholar. For the 77/78 sample, the average productivity measure was 1.55 for publications and 1.15 for accounting publications only.

Suggested use. This study may prove useful in developing benchmarks for assessing faculty research productivity. The results can be used as benchmarks for new scholars publishing in academic, noneducation journals. The results can be used as benchmarks for broader productivity that include education and practitioner publications along with other academic

publications. The results should prove helpful to decision makers, whether they are senior faculty evaluating the research productivity of junior faculty or junior faculty evaluating themselves.

Fogarty (2004)

Identification of weaknesses of previous published benchmarks. Previous studies failed to provide an adequate picture of the research activity of more senior accounting faculty. This paper explores the effects of institutional and individual antecedents on the publication productivity of accounting academics.

Basis of benchmarks. Based on the activity level of 1,384 faculty who earned their doctoral degree more than twenty-five ago over a ten-year period at the annual meetings of the American Accounting Association (AAA), not on journal publications.

Method. The contribution of each individual faculty to the ten annual meetings was measured by total number of meetings in which the individual participated as an author of research and by the number of each instance of research participation by each faculty member.

Summary of benchmarks. Accounting faculty who received their doctoral degrees from high status universities are more likely to be sustained accounting researchers. Faculty possessing current appointments at top schools are likely to continue to do more research well into the course of their careers. Higher levels of early career research activities are associated with the continuation of research.

Suggested use. Faculty in the seasoned cohort should have a personal interest in benchmarking their levels of productivity. Younger faculty should also take note since it will not be long until they move into that segment of their careers. Administrators of accounting departments and others involved in the process of supervising accounting faculty should benefit from the analysis as a means of gauging expectations, including those who work in the accreditation effort

Everett, Klamm and Stoltzfus (2004)

Identification of weaknesses of previous published benchmarks. Prior studies examined quantity of publications, quality of publications, or a combination of quantity or quality. They do not consider the breadth and depth of the aggregate publication productivity of faculty.

Basis of benchmarks. Summarizes the publication productivity of 87 accounting doctoral programs in two subsets of the 30 highest-rated academic accounting journals over five years (1992–1996).

Method. Compiled lists of 30 journals using previous studies of journal rankings. Publication data of individually named faculty for the five-year period, 1992–1996, were collected for all tenured and tenure-track faculty members at 87 universities for the 1996–1997 academic year. *Quality* was determined by dividing a cumulative quality score for a university by the number of tenured and tenure-track faculty at the university in the fifth year of the study (the 1996–1997 academic year). *Breadth* was determined by dividing the number of tenured and tenure-track faculty who published at least one paper within the five-year period by the number of tenured and tenure-track faculty at the university in the fifth year of the study. *Depth* was based on a ratio between the number of faculty publishing in a given five-year period and the number of Ph.D. graduates within the same five-year period.

Summary of benchmarks. Quality: A median quality score of 0.79 for the top 30 journals and 0.69 for the top 15 journals. All universities in the study had at least one publication in the five-year window when 30 journals were examined. All universities except one had at least one publication in the five-year period when 15 journals were examined, and the highest score for this analysis was 4.22. *Breadth:* The median percentage of faculty publishing in the top 30 journals was 52%, while the median percentage for the top 15 journals was 38%. The highest possible score (100%) was attained by three universities and this score was unchanged when the number of journals was decreased from 30 to 15. *Depth:* a median of 0.75 for the subset of 30 journals and a median

depth measure of 0.57 for the subset of 15 journals. The highest five-year score for both the 30-journal and the 15-journal results was 6.00.

Suggested use. Aggregating the results by a university should provide useful benchmarks for administrators of doctoral programs in accounting to assess their program's productivity and to compare their results with other universities.

Glover, Prawitt and Wood (2006)

Identification of weaknesses of previous published benchmarks. There is a limited amount of current, relevant publication data that can be used by faculty and administrators to plan and evaluate research productivity and to set research criteria. Previous studies examining faculty research productivity do not present a clear picture of the publication records of successful candidates at the time of promotion.

Basis of benchmarks. Sample includes approximately 20 percent of all faculty promotions at all four-year institutions with accounting programs in the U.S. Sample included 156 professors promoted to associate and 85 professors promoted to full professor.

Method. Documents the publication data of accounting faculty promoted from 1995 to 2003 at the top 75 research-oriented accounting programs. Focuses on schools that publish the most in top-tier journals. Ranked accounting program based on the total number of pages published in 20 "top-tier" business research journals.

Summary of benchmarks. Seventy-six point three percent of professors promoted to associate professor had published at least one article in the Top 3 accounting journals. Forty-six point eight percent of these professors had published at least two articles in these top journals. At the time of promotion to full professor, every professor in the study had published at least one article in a journal ranked in the Top 15 journals. More than 15 percent of the faculty

promoted to full professor had published 10 or more articles in the Top Business category. Almost 30 percent had published 10 or more articles in the Top 6 category.

Suggested use. The study should be helpful to faculty members and administrators at the top 75 accounting research programs, as well as other schools with similar scholarship targets, in setting goals, in establishing criteria, and in evaluating performance. This study will contribute to the quality and consistency of tenure and promotion discussions and decisions by making available descriptive data on the publication records of successful candidates. Decision makers can use data in developing or evaluating research expectations for promotion decisions.

Reinstein and Calderon (2006)

Identification of weaknesses of previous published benchmarks. The literature does not indicate the extent to which accounting departments rely on published or self-generated journal rankings in evaluating faculty research and scholarship.

Basis of benchmarks. Surveyed 145 members of the AAA's Accounting Leadership Program Group.

Method. Survey ranked journals that accounting departments used for promotion, tenure, merit pay and other purposes. Used accounting department journal rankings to produce composite rankings for all listed accounting journals.

Summary of benchmarks. Accounting departments use journal rankings as a means of reducing the complexity of their processes for evaluating faculty scholarship. There was a statistically significant difference between the number of faculty in departments that use journal rankings and those that do not such rankings. Departments with doctoral programs and separate AACSB accreditation are more likely than other departments to use journal rankings. Journal rankings are most heavily concentrated among non-elite Ph.D.-granting programs. Ten responding Ph.D. programs use a ranked publications list of an average of 39 journals to assess faculty productivity.

Suggested use. Accounting faculty and administrators assign much importance to scholarship in making merit, tenure and promotion decisions in accounting programs. Academic administrators often seek objective research productivity data to make performance evaluation, hiring, tenure and promotion decisions. Both administrators and faculty are particularly interested in benchmark data to help set research productivity standards, to measure their own progress and to assess the progress of others.

Kerr, Simkin and Mason (2009)

Identification of weaknesses of previous published benchmarks. The most popular measure of “research productivity” has been counts of papers published in “premier” journals. Accounting faculty often publish in specialization journals that may not appear in the journals on the “count lists.”

Basis of benchmarks. Survey by mail of faculty of accounting departments of AACSB accredited universities.

Method. Surveyed faculty about the total number of blind-refereed articles they had published to date by their area of research specialization.

Summary of benchmarks. Overall mean publication rate was 15.4 per faculty member articles regardless of rank, years of service, or gender. Specialization areas of auditing and tax produced the highest publication rates, each with 17.0 publications per faculty member. Management accounting had the lowest publication level of 11.5 publications faculty member.

Suggested use. The results provide several benchmarks for accounting faculty. One useful set of values is the total number of publications, by rank, for accounting faculty—i.e., about four articles for assistant professors, 11.6 articles for associate professors, and 20 articles for full professors. A second benchmark is the overall publication rate of almost exactly one blind-refereed journal

publication per year. Provides a reasonable target for accounting faculty at AACSB-accredited schools.

Walker, Fleischman and Stephenson (2010)

Identification of weaknesses of previous published benchmarks. Only a few studies examine perceptual differences between faculty and administrators regarding tenure and promotion criteria. In addition to a ranked journal list, accounting departments should also consider establishing benchmarks for the quantity of publications necessary for tenure and promotion that incorporate appropriate quality category weightings.

Basis of benchmarks. Three email surveys in 2006, 2007, and 2010 of counting department administrators of Ph.D. and non-Ph.D. accounting programs, and AACSB and non-AACSB accredited institutions.

Method. Surveys asked for information about department-specific written tenure and promotion standards.

Summary of benchmarks. The first study that reports on accounting-specific written standards for accounting tenure and promotion purposes. Second study to review specific accounting documented journal lists that doctoral and non-doctoral AACSB-accredited institutions actually use for purposes of tenure and promotion deliberations. Eighty percent of schools in sample do not have documented tenure and promotion standards for research. Sixty-three percent of non-doctoral schools do not have documented tenure and promotion standards for research. Ninety-six percent of non-AACSB-accredited schools do not have documented tenure and promotion standards for research. Nine percent used ranked journal lists as part of faculty evaluations.

Suggested use. May provide a starting point for departments of accounting that wish to initiate written tenure and promotion policies and faculty workload assignments based on research

productivity. An accounting department may wish to adapt a ranked journal list so that it is consistent with its unique mission.

Hasselback, Reinstein and Abdolmohammadi (2012)

Identification of weaknesses of previous published benchmarks. Increasing attention to faculty research productivity suggests a need for reliable benchmarks.

Basis of benchmarks. Based on records of 5,607 accounting doctoral graduates from 1971–2005. Research productivity is measured using unadjusted number of published articles (Best 3, Best 13, Best 24, and Best 40 journals); published articles adjusted for journal quality scores; published articles adjusted for co-authorship; and published articles adjusted for both journal quality and co-authorship. Benchmarks based on faculty productivity in four sets of journals from 1971–2005.

Method. Based on publication records of 5,607 accounting doctoral graduates between 1971 and 2005 who published 22,579 articles in the “best” 40 journals through 2009.

Summary of benchmarks. Average publication productivity of accounting faculty per year had steadily increased over the previous 35 years under study. They classified journals in four sets—Top 3, Next 10, Next 11 and Next 16—from 1971–2005 and for each year of 2001–2005. On average each faculty member published unadjusted 0.20 articles per year, and 0.28 after adjusting for journal quality.

Suggested use. The benchmarks identified in this study can help with tenure, promotion, merit pay, appointment and renewal of chaired professorships, and other resource allocation decisions. Focusing on individual research benchmarks can help identify (1) the research productivity of faculty members’ national peers; (2) criteria for awarding new faculty members ranks of associate professor or full professor, or tenure; and (3) standards to select or retain chaired professors.

Walker, Fleishman and Stephenson (2013)

Identification of weaknesses of previous published benchmarks. Walker, Fleishman and Stephenson's (2013) benchmarks are not benchmarks in the same sense as the previous eleven. That is, the benchmarks are based on a point system rather than a comparison with publications by other faculty. Walker, Fleishman and Stephenson report on the University of Wyoming's Department of Accounting's experiences in developing a research productivity policy for use in tenure and promotion and workload assignment decisions. Motivations for developing a research productivity policy included uncertainty of accounting faculty concerning research requirements, expressed department policy and college standards were vague, and differences in the research "landscape" may lead to disagreement among college administrators about the significance of research productivity in accounting. Evidence suggests that the literature on journal quality and publication output of accounting faculty and related ranked journal lists is not applied formally for tenure and promotion purposes on a wide scale in accounting.

Basis of benchmarks. Attempts to capture realistic, relative quality differences using a weighting procedure based on perceived publication difficulty representing the amount of effort.

Method. Awards points for publications in journals based on perceived publication difficulty.

Summary of benchmarks. Six points are awarded for publications in A+ journals, 4 points awarded for publications in A journals, 2 points awarded for publications in A- journals, 1 point awarded for publications in B+ journals, and 0.5 points awarded for publications in all other journals.

Suggested use. Faculty and administrators can use the research findings and the model as a starting point to develop their own policies consistent with their specific departments' mission.

SUMMARY

First, contrary to their assertions, it can easily be seen that the published benchmarks do not identify a set of objective standards but are based on perceptions and thus are not true benchmarks. To label them as benchmarks thus is inaccurate.

Second, literally every published benchmark of the last 20 years begins with the quantity of publications of faculty members which is then augmented by perceived journal rankings, doctoral program rankings, institution rankings, or individually named faculty. Variants include expanding the quantity of publications and journal rankings by rankings of institutions, programs, or persons.

Published benchmarks rely on surveys of perceptions of journal “quality” or program rankings, the choice of which set of published benchmarks best measures faculty research productivity will differ among administrators based on their individual perceptions which diminishes the usefulness and acceptance of published benchmarks. Furthermore, each succeeding published benchmark purports to be an improvement over previously published benchmarks by identifying weaknesses of previously published benchmarks and expanding the number or types of journals, or the number or types of programs, which suggests that the use of previously published benchmarks should be abandoned and the latest set of published benchmarks should be adopted.

Our study thus fills the gap between the published benchmarks and their actual use by administrators by attempting to determine whether administrators are familiar with the published benchmarks and if so, whether the published benchmarks are actually used by administrators. The value of published benchmarks will be validated by their usage in actual practice. Their non-use, however, will confirm the existence of a disconnect between published benchmarks and the actual use of published benchmarks.

Hypotheses and Methodology

Hypotheses

There are four hypotheses. Researchers who have developed faculty research productivity benchmarks suggest that their benchmarks can be used for purposes of hiring, tenure, promotion, or merit pay decisions. Therefore, the basic research hypothesis is that administrators do not use published benchmarks to evaluate faculty research productivity for purposes of hiring, tenure, promotion, or merit pay decisions. This is broken down into one hypothesis for each type of decisions:

HYPOTHESIS 1a. *Administrators use published benchmarks to evaluate faculty research productivity for hiring decisions.*

HYPOTHESIS 2a. *Administrators use published benchmarks to evaluate faculty research productivity for promotion decisions.*

HYPOTHESIS 3a. *Administrators use published benchmarks to evaluate faculty research productivity for tenure decisions.*

HYPOTHESIS 4a. *Administrators use published benchmarks to evaluate faculty research productivity for merit pay decisions.*

Methodology

The survey was hosted on SurveyMonkey and invitations were sent via email to 1,435 deans, chairs, and heads of accounting departments email addresses listed in Hasselback's (2016) *Accounting Faculty Directory*. The survey asked whether they are familiar with published accounting research productivity benchmarks and if so, whether they used the benchmarks to evaluate the accounting research productivity of their faculty or what decisions they may make based on the benchmarks if the benchmarks were used.

The invitations to participate were sent on Feb, 1, 2017 and respondents were given four weeks to begin and complete the survey. Respondents had to complete the survey in one sitting in order to discourage attempts to search for the published benchmarks while completing the survey. Three reminders were sent over the following four weeks.

The survey consisted of ten background questions and three questions for each of the 12 benchmarks. (Appendix B). Respondents were asked to indicate their position (chair, dean, etc.), whether their institutions were AACSB accredited, what is the highest degree awarded by their institution, whether their institution was public or private, sectarian/religious, and whether classified as a research institution. They were then asked if they were familiar with the published benchmarks indicated. If they were familiar with particular published benchmarks, they were then asked whether they used the published benchmarks and if so, whether they used the benchmarks to evaluate faculty for hiring, tenure, promotion, or merit pay decisions. If they were not familiar with particular published benchmarks, they were directed to the next question. Thus, not all respondents answered the same questions.

RESULTS AND ANALYSIS

RAW DATA AND DESCRIPTIVE STATISTICS

The raw data and descriptive statistics are given in Table 1. The panels in Table 1 present the percentages of respondents who are familiar with and who actually use each benchmark by demographics (position, highest degree, time in position, AACSB accredited, sectarian/religious, research institution, and public/private).

[Insert Table 1, Panels A, B, and C about here]

Of the 1,435 email invitations sent to the population of accounting administrators, 234 were returned for invalid email addresses leaving 1,201. After eliminating incomplete responses, 241 completed all questions for a 20.1% completion rate. The high number and response rate suggest that the issue of benchmarks is considered important by accounting department administrators.

Forty-six (19.09%) were Deans of Business or Accounting Schools, 129 (57.68%) were Chairs of Accounting Departments, 46 (19.09%) were Directors or Heads of Accounting Programs, and 23 (9.54%) were Other. Twenty-three (9.54%) were in their positions less than 1 year, 127 (52.70%) 1-5 years, 61 (25.31%) 6-10 years, 24 (9.96%) 11-20 years, and 6 (2.49%) more than 20 years.

One hundred ninety-one (79.25%) were AACSB accredited and 92 (38.17%) were AACSB Accounting accredited. One hundred forty-six were public and 95 (39.42%) were private. Fifty-five were considered as sectarian or religious, while 112 (46.47%) were classified as research institutions. The highest degree awarded was a doctorate degree 83 (34.44%), 111(46.06%) awarded a master's degree, and 47 (19.50%) awarded bachelor's degrees.

In chronological order according to publication date, 116 (48.13%) were familiar with Hasselback and Reinstein's (1995a) benchmarks, but only 8 (6.90%) of those familiar with the benchmarks actually used them, or (3.32%) of the total number of respondents. Eight used the benchmarks for tenure and promotion decisions, six for merit pay decisions, and one for hiring decisions.

For Hasselback and Reinstein (1995b), 108 (44.81%) were familiar with the benchmarks, but only 7 (6.48%) of those familiar with the benchmarks, 2.90% of the total, actually used them. Seven used the benchmarks for tenure and promotion decisions, four for merit pay decisions, and one for hiring decisions.

One hundred (41.49%) were familiar with Hasselback, Reinstein and Schwan's (2000) benchmark, but only 4 (4.00%) used them, which represented 1.66% of the total. Three used the benchmarks for tenure and promotion decisions, and one for merit pay decisions.

For Christensen, Finger and Latham (2002), 83 (34.44%) were familiar with the benchmarks, but none of the respondents used them. While this is not the lowest number of those familiar with benchmarks, it is the lowest number and percent of those familiar with the benchmarks who use them, as well as the lowest number and percent of total respondents.

Seventy-one (29.46%) were familiar with Fogarty's (2004) benchmarks, but only 2 (2.82%, 0.83% of total respondents), actually used them. Two used the benchmarks for tenure, promotion, and hiring decisions, one used them for merit pay decisions.

For Everett, Klamm and Stoltzfus (2004), 71 (29.46%) were familiar with the benchmarks, but only 1 (1.41%) used them representing 0.41% of the total for tenure and promotion decisions.

Out of 106 (43.98%) familiar with Glover, Prawitt and Woods (2006) benchmarks, 18 (16.98%, 7.47% of total) used them. This represents the highest number and highest percent of those familiar with the benchmarks who use them, as well as the highest number and highest percent of the total. Thirteen used the benchmarks for tenure and promotion decisions, four for merit pay decisions, and five for hiring decisions.

Seventy-eight (32.37%) were familiar with Reinstein and Calderon's (2006) benchmarks, but only 5 (6.41%) actually used them, representing 2.07% of the total. Five used the benchmarks for tenure and promotion decisions, and two used them for merit pay and hiring decisions.

Of the 241 total respondents, 63 (26.14%) were familiar with Kerr, Simkin and Mason's (2009) benchmarks, which is 1.24% of the total. One used the benchmarks for tenure, promotion, and hiring decisions, two used them for merit pay decisions.

Sixty-eight were familiar with Walker, Fleischman and Stephenson's (2010) benchmarks, but only one (1.47%) used them representing 0.41% of the total. Thirteen used the benchmarks for tenure and promotion decisions, four for merit pay decisions, and five for hiring decisions. The benchmarks were used for merit pay decisions.

Eighty-eight (36.51%) were familiar with Hasselback, Reinstein and Abdolmohammadi's (2012) benchmarks. Five (5.68%) of those familiar with the benchmarks (2.07% of the total) used them. Eight used the benchmarks for tenure decisions, seven for promotion decisions, five for merit pay decisions, and one for hiring decisions.

Finally, sixty-four (26.56%) were familiar with Walker, Fleishman and Stephenson's (2013) benchmarks, but only two (3.13%) used them, representing 0.83% of the total. One used them for tenure, promotion, and merit pay decisions.

There were several comments in the "Other" category to the effect that "we use our own benchmarks." This lends support to our hypotheses that overall, published benchmarks are at best of questionable usefulness.

ANALYSIS

The results were analyzed using chi-square. The analysis is given in Table 2. The panels in Table 2 present the chi-square results of demographic variables, analysis by familiarity and use, and analysis by decisions.

All null hypotheses are rejected at the .000 level. Administrators do not use published accounting research productivity benchmarks to evaluate faculty research productivity for hiring, tenure, promotion, and merit pay decisions which basically means that while published benchmarks have contributed to the literature, they have not contributed to the evaluation of faculty research productivity as proposed by the researchers.

Familiarity and use by demographics

Data analysis. Total variables for familiarity and use were coded. If respondents indicated they were familiar with any of the 12 benchmarks, Total-familiar was coded as 1 (yes they are familiar with the benchmarks) or 0 (no they are not familiar). If respondents indicated they used any of the 12 benchmarks, Total-use was coded as 1 (yes they use the benchmarks) or 0 (no they do not use the benchmarks). Crosstabs were then calculated for each demographic variable by Total-familiar and Total-use. A Chi-square test of independence was calculated for each of the crosstabs.

[Insert Table 2, Panels A, B, and C about here]

Position. There was a significant difference in familiarity with published benchmarks according to position, $\chi^2 (3, N=241) = 19.46, p = .000$. The highest percentages who indicated they are familiar with published benchmarks were Director/Head of Accounting (78.8%) and Chair of Accounting Department (66.2%). The lowest percentage who reported they are familiar with published benchmarks was Dean of Business/Accounting School (21.45%).

There was no significant difference in use of published benchmarks among the positions, $\chi^2 (3, N=241) = 5.15, p = .161$. The percentage of those who indicated they use the published benchmarks over all positions was 13.7%.

Years in present position. There was no significant difference in familiarity with published benchmarks based on years of experience, $\chi^2 (4, N=241) = 4.82, p = .306$. The percentage of those who reported they are familiar with published benchmarks over all the years of experience was 60.6%.

There was no significant difference in use of published benchmarks based on years of experience, $\chi^2 (3, N=241) = 7.50, p = .101$. The percentage who indicated they use published benchmarks over all years of experience was for all the positions was 13.7%.

Highest degree of institution. There was a significant difference in familiarity with published benchmarks based on the highest degree awarded by the institution, $\chi^2 (2, N=241) = 48.51, p = .000$. The highest percentage of familiarity with published benchmarks were from institutions where the highest degree is the doctorate (89.2%) followed by those institutions where the highest degree is the masters (51.4%) and lastly institutions where the highest degree is Bachelors (31.9%).

There was also a significant difference in use of published benchmarks based on the highest degree awarded by the institution, $\chi^2 (2, N=241) = 9.40, p = .009$. The highest percentage of use of published benchmarks were at institutions where the highest degrees awarded is a doctorate

(18.1%) and masters (16.2%). None of those in institutions where bachelors is the highest degree indicated they use the benchmarks.

AACSB accredited. There was a significant difference in both familiarity, $\chi^2 (1, N=241) = 9.34, p = .002$, and use, $\chi^2 (1, N=241) = 10.50, p = .001$, of published benchmarks according to whether an institution was AACSB accredited. The highest percentage of those who indicated they were familiar with published benchmarks were from AACSB accredited schools (72.8%). The lowest percentage were from the non-AACSB accredited schools (53.0%). The highest percentage who indicated they use published benchmarks were from AACSB accredited schools (22.8%). The lowest percentage who indicated they use published benchmarks were from non-AACSB accredited schools (8.1%).

Public or private. There was no significant difference in familiarity with published benchmarks, $\chi^2 (1, N=241) = 3.2, p = .077$, based on whether the institution was public or private. The overall percentage of those familiar with published benchmarks was 60.6%.

There was, however, a significant difference in use of published benchmarks based on, $\chi^2 (1, N=241) = 5.31, p = .021$, based on whether the institution was public or private. The highest percentage who indicated they use published benchmarks were from public schools (17.8%). The lowest percentage were from private schools (7.4%).

Sectarian/religious. There was no significant difference in familiarity with published benchmarks, $\chi^2 (1, N=241) = 2.79, p = .095$, according to whether or not the institution was sectarian or religious. The overall percentage of those familiar was 60.6%

There was a significant difference in use of published benchmarks, $\chi^2 (1, N=241) = 4.09, p = .043$, according to whether or not the institution was sectarian or religious with the highest percentage of those who indicated they use published benchmarks from non-sectarian/religious schools (16.1%), and the lowest percentage from sectarian/religious schools (5.5%).

Research institution. There was a significant difference in familiarity with published benchmarks, $\chi^2 (1, N=241) = 47.46, p = .000$ based on whether the institution was classified as a research institution. The highest percentage who indicated they were familiar with published benchmarks were from research institutions (63.9%). The lowest percentage were from non-research institutions (40.3%).

There was a significant difference in use of published benchmarks, $\chi^2 (1, N=241) = 13.18, p = .000$, based on whether the institution was classified as a research institution. The highest percentage who indicated they published the benchmarks were from research institutions (22.3%) and the lowest percentage were from non-research (6.2%).

Summary. Higher percentages of chairs and directors of accounting departments indicated they were familiar with published benchmarks than the percentage of deans. About the same percentages of the levels of years of experience indicated they are familiar with published benchmarks.

Higher percentages of respondent at institutions where a doctorate was the highest degree awarded indicated they are familiar with published benchmarks. Higher percentages of respondents at AACSB accredited institutions, public institutions, non-sectarian/religious institutions, and research institutions indicated they are familiar with published benchmarks.

Higher percentages of chairs and directors of accounting departments also indicated they use published benchmarks than the percentage of deans. Those with 1-5 years of experience in their present position were the highest percentage of those who use published benchmarks.

Approximately the same percentages of respondent at institutions where a doctorate was the highest degree awarded indicated they are familiar with published benchmarks. Higher percentages of respondents at AACSB accredited institutions, public institutions, non-

sectarian/religious institutions, and research institutions indicated they use published benchmarks.

Familiarity and Use of Published Benchmarks

Frequencies of “yes” responses for familiarity and use were calculated for each benchmark. A nonparametric one-sample chi-square was used to test for differences in the response rates.

Hasselback and Reinstein (1995a). There was no significant difference ($p = .562$) in the percentage of those who are familiar with (48.1%) and those who are not familiar with (51.9%) this benchmark. About the same percentage of the respondents were familiar as those who were not familiar with this benchmark. There was a significant difference ($p = .000$) in the percentage of those who indicated they use (3.3%) and those who do not use (96.7%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Hasselback and Reinstein (1995b). There was no significant difference ($p = .107$) in the percentage of those who are familiar with (44.8%) and those who are not familiar with (55.2%) this benchmark. There was a significant difference ($p = .000$) in the percentage of those who indicated they use (2.9%) and those who do not use (97.1%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Hasselback, Reinstein and Schwan (2000). There was a significant difference ($p = .008$) in the percentage of those who are familiar with (41.5%) and those who are not familiar with (58.5%) this benchmark. A significantly lower percentage of the respondents are familiar with this benchmark. There was also a significant difference ($p = .000$) in the percentage of those who indicated they use (1.7%) and those who do not use (98.3%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Christensen, Finger and Latham (2002). There was a significant difference ($p = .000$) in the percentage of those who are familiar with (34.4%) and those who are not familiar with (65.6%)

this benchmark. A significantly lower percentage of the respondents are familiar with this benchmark. There was a significant difference in percentage of those who indicated they use (0.0%) and those who do not use (100.0%) this benchmark. None of the respondents use this benchmark.

Fogarty (2004). There was a significant difference ($p = .000$) in the percentage of those who are familiar with (29.5%) and those who are not familiar with (70.5%) this benchmark. A significantly lower percentage of the respondents are familiar with this benchmark. There was also a significant difference ($p = .000$) in the percentage of those who indicated they use (2.8%) and those who do not use (97.2%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Everett, Klamm and Stoltzfus (2004). There was a significant difference ($p = .000$) in the percentage of those who are familiar with (29.5%) and those who are not familiar with (70.5%) this benchmark. A significantly lower percentage of the respondents are familiar with this benchmark. There was a significant difference ($p = .000$) in percentage of those who indicated they use (0.4%) and those who do not use (99.6%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Glover, Prawitt and Wood (2006). There was no significant difference ($p = .062$) in the percentage of those who are familiar with (44.0%) and those who are not familiar with (56.0%) this benchmark. About the same percentage of the respondents were familiar as those not familiar with this benchmark. There was a significant difference ($p = .000$) in the percentage of those who indicated they use (7.5%) and those who do not use (92.5%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Reinstein and Calderon (2006). There was a significant difference ($p = .000$) in the percentage of those who are familiar with (32.4%) and those who are not familiar with (67.6%) this benchmark. A significantly lower percentage of the respondents are familiar with this benchmark. There was

a significant difference ($p = .000$) in percentage of those who indicated they use (2.1%) and those who do not use (97.9%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Kerr, Simkin and Mason (2009). There was a significant difference ($p = .000$) in the percentage of those who are familiar with (26.1%) and those who are not familiar with (73.9%) this benchmark. There was also a significant difference ($p = .000$) in percentage of those who indicated they use (.8%) and those who do not use (99.2%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Walker, Fleischman and Stephenson (2010). There was a significant difference ($p = .000$) in the percentage of those who are familiar with (28.2%) and those who are not familiar with (71.8%) this benchmark. A significantly lower percentage of the respondents are familiar with this benchmark. There was a significant difference ($p = .000$) in the percentage of those who indicated they use (0.0%) and those who do not use (99.6%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Hasselback, Reinstein, Abdolmohammadi (2012). There was a significant difference ($p = .000$) in the percentage of those who are familiar with (38.5%) and those who are not familiar with (61.5%) this benchmark. A significantly lower percentage of the respondents are familiar with this benchmark. There was also a significant difference ($p = .000$) in the percentage of those who indicated they use (2.1%) and those who do not use (97.9%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Walker, Fleishman and Stephenson (2013). There was a significant difference ($p = .000$) in the percentage of those who are familiar with (26.6%) and those who are not familiar with (73.4%) this benchmark. A significantly lower percentage of the respondents are familiar with this benchmark. There was a significant difference ($p = .000$) in the percentage of those who indicated

they use (0.8%) and those who do not use (99.2%) this benchmark. A significantly lower percentage of the respondents actually use this benchmark.

Summary. The percentages of respondents who are familiar with the benchmarks ranged from 26.1% to 44.8% indicating that less than half of the respondents indicated they are familiar with the benchmarks. The percentages of respondents who actually use the benchmarks ranged from 0.4% to 7.5% revealing that very low percentages of the respondents indicated they actually use the benchmarks.

Test of hypotheses

All hypotheses are rejected at the .000 level of significance for all decisions. Administrators do not use published benchmarks to make hiring, promotion, tenure, or merit pay decisions.

Response frequencies were calculated for each of the four decisions (tenure, promotion, merit pay, hiring) by each benchmark. Non-parametric one-sample chi-square tests were used to determine if there was a significant difference in the proportions who use the benchmark and don't use the benchmark for faculty research productivity decisions. For some of the benchmarks the respondents indicated they did not use the benchmark for decisions, thus these variables were constant and chi-square could not be calculated.

HYPOTHESIS 1. Administrators use published benchmarks to evaluate faculty research productivity for hiring decisions.

Hypothesis 1 was rejected at the .000 level of significance. Administrators do not use the published benchmarks to evaluate faculty research productivity for hiring decisions. The percentages of the respondents who actually use the benchmarks for evaluating faculty research productivity for hiring decision ranged from 0% to 5.4%. Two of the benchmarks were not used at all (Christensen, Finger and Latham (2002), and Walker, Fleischman and Stephenson (2010))

HYPOTHESIS 2. Administrators use published benchmarks to evaluate faculty research productivity for promotion decisions.

Hypothesis 2 was also rejected at the .000 level of significance. administrators do not use the published benchmarks to evaluate faculty research productivity for promotion decisions. The percentages of the respondents who actually use the benchmarks for evaluating faculty research productivity for promotion decision ranged from 0% to 5.4%. Again, two of the benchmarks were not used at all (Christensen, Finger and Latham (2002), and Walker, Fleischman and Stephenson (2010))

HYPOTHESIS 3. Administrators use published benchmarks to evaluate faculty research productivity for tenure decisions.

This hypothesis was rejected at the .000 level of significance. Administrators do not use the published benchmarks to evaluate faculty research productivity for tenure decisions. The percentages of the respondents who actually use the benchmarks for evaluating faculty research productivity for tenure decisions ranged from 0% to 2.5%. Three of the benchmarks were not used at all (Christensen, Finger and Latham (2002), Everett, Klamm and Stoltzfus (2004), and Walker, Fleischman and Stephenson (2010))

HYPOTHESIS 4. Administrators use published benchmarks to evaluate faculty research productivity for merit pay decisions.

Finally, Hypothesis 4 was rejected at the .000 level of significance. Administrators do not use the published benchmarks to evaluate faculty research productivity for merit pay decisions. The percentages of the respondents who actually use the benchmarks for evaluating faculty research productivity for hiring decision ranged from 0% to 2.1%. Five of the benchmarks were not used at all (Hasselback, Reinstein and Schwan (2000), Christensen, Finger and Latham (2002), Everett, Klamm and Stoltzfus (2004), and Walker, Fleischman and Stephenson (2010), and Walker, Fleishman and Stephenson (2013))

SUMMARY

Very small percentages of respondents indicated they actually use the benchmarks (0.0% to 5.4%). The percentages were significantly different from the percentages of those who did not indicate they used the benchmark for decision. Two of the benchmarks were not selected as being used for any of the four decisions (Christensen, Finger and Latham (2002) and Walker, Fleischman and Stephenson (2010)). Glover, Prawitt and Wood (2006) had the highest percentages of actual use (1.7% to 5.4%). The percentage of respondents who indicated they use the benchmarks for hiring decisions was generally the lowest for all the benchmarks.

CONCLUSION

The evidence does not support the claims that published benchmarks are useful to administrators in hiring, tenure, promotion, and merit pay decisions. As the results of this study show, none of the benchmarks published in the last 20 years have been widely adopted. One would expect that as the importance evaluating faculty research productivity has increased that the use of published benchmarks would likewise increase. But that is not the case indicating that the published benchmarks have little practical use.

Publishing research productivity benchmarks seems to be an industry unto itself. While the published benchmarks include the obligatory “contribution to the literature,” there seems to be no contribution to practical use. The greatest use of published benchmarks studies is to provide citations in subsequent benchmark studies. The benchmarks are not used by administrators to evaluate faculty research productivity.

While quantity of publications alone reveals little about faculty research productivity, the value of adding variants to quantity of publications such as the subjective ranking of institutions, programs, or persons is doubtful. Such rankings serve more to enhance the reputations of particular institutions, programs, or persons rather than to provide actual benchmarks of faculty research productivity. Such benchmarks are misguided and have little relationship to actual faculty research productivity.

Evaluating faculty research productivity based on subjective perceptions of particular journals, institutions, programs, or persons rather than on standard points of reference tends to skew research productivity results away from an objective baseline toward a particular journal, institution, program, or person. As noted by Stone (2017), what is observed in practice is “journal rankings based upon surveys of the opinions of scholars, with seemingly little regard for objective, verifiable evidence about the processes of the journals or the quality of their output.” Journal rankings are merely, according to Stone, “popularity contests.”

The results of this study suggest that continued attempts to develop benchmarks for evaluating accounting faculty research productivity benchmarks using institution, journal, or program rankings is the equivalent of hunting a snark⁶—an exercise in futility. Furthermore, accounting research productivity benchmarks are a moving target, subject to changes in the subjective rankings of journals, programs, or universities rankings which make realistic comparisons next to impossible.

Establishing future benchmarks must be different if they are to be useful. For example, establishing baseline benchmarks for research publication activity would be an important first step. Considering contributions to things other than the literature, such as contributing to social well-being or the improvement of laws or regulations, could also be included. To date, none of the published benchmarks attempt to establish baselines other than the system used by the University of Wyoming (Walker, Fleishman, and Stephenson, 2013) which, while not a true benchmark since it awards points for publications in certain subjectively ranked journals, at least make faculty aware of the point of reference and standard by which to measure their research productivity is to be judged.

This study did not attempt to discover whether accounting faculty use published benchmarks to compare their own research productivity with that of other accounting faculty as suggested

⁶ Hunting a snark refers to a poem published in 1876 by Charles Lutwidge Dodgson (aka Lewis Carroll, who wrote “Alice in Wonderland” and “Jabberwocky”) about hunting a non-existent, mythical creature.

Hasselback, Reinstein and Schwan (2003). But such comparisons would yield little useful information anyway. For one faculty member to compare his or her research productivity with that of another faculty member transforms academic research into a rivalry (Huber 2015; Stone 2017), a competition where one will attempt to exceed the output of another. Such productivity statistics are akin to those of football players where exceeding the number of passing yards or receiving yards of veteran players becomes the goal of every rookie right out of college⁷ as can be seen in comparing the “senior cohort” (Fogarty 2004) with “new accounting scholars’ publications” (Christensen, Finger, and Latham 2002). Benchmarks that merely compare faculty research productivity with the research productivity of other faculty cannot legitimately be used to evaluate faculty research productivity. Not only are such comparisons not fair to faculty because they are constantly changing, but they misdirect faculty research efforts into duplicating the methodology used in previous research in an effort to have their own research published, particularly in elite journals. (Huber 2015, 2016a, 2016b).

Nor did the study attempt to determine what benchmarks or other measures administrators used to evaluate faculty research productivity. The study was limited only published benchmarks and whether administrators used them as suggested by the benchmark studies.

Future research into developing legitimate and useful faculty research productivity benchmarks is necessary, but not by merely adding more or different journal or program rankings which. Faculty research productivity must be measured by true benchmarks, not by standards that change according to the whim of administrators’ journal or program rankings.

Future research into how, or whether, published accounting faculty productivity benchmarks are used by university administrators in other countries could also provide useful insights into how to construct benchmarks that can be useful to accounting administrators in U.S. universities.

⁷ It is interesting to note that the American Accounting Association sponsors an annual “Ph.D. Rookie camp.” See, e.g., “2017 Accounting PhD Rookie Recruiting and Research Camp,” <http://aaahq.org/Meetings/2017/RookieCamp>.

References

- Association for the Advancement of Collegiate Schools of Business (AACSB). 2003. *Eligibility procedures and standards for business accreditation*. St. Louis, MO: AACSB.
- Bernardi. R.A., K.M. Zamojcin, and T.L. Delande. 2016. Ranking accounting authors and departments in accounting education: different methodologies – significantly different results. *Accounting Education: An International Journal*: 25(6), 568-597.
- Chan, K.C., C.R. Chen, and L.T.W. Cheng. 2005. Ranking research productivity in accounting for Asia-Pacific Universities. *Review of Quantitative Finance and Accounting*, 24: 47-64.
- Christensen, A.L., C.A. Finger, and C.K. Latham. 2002. New accounting scholars' publications in accounting and nonaccounting journals. *Issues in Accounting Education* 17(3): 234-251.
- Everett, J.O., B. Klamm, and R. Stoltzfus. 2004. Developing benchmarks for evaluating publication records at doctoral programs in accounting. *Journal of Accounting Education* 22(3): 229–252.
- Fogarty, T.J. 2004. Sustained research productivity in accounting: A study of the senior cohort. *Global Perspectives on Accounting Education* 1: 31-58.
- Glover, S.M., D.F. Prawitt, and D.A. Wood. 2006. Publication records of faculty promoted at the top 75 accounting research programs. *Issues in Accounting Education* 21(3): 195-218
- Hasselback, J.R. 2016. *Accounting faculty directory*. Englewood Cliffs, NJ: Prentice-Hall.
- Hasselback, J.R., and A. Reinstein. 1995a. Assessing accounting doctoral programs by their graduates' research productivity. *Advances in Accounting* II: 61-86.
- Hasselback, J.R., and A. Reinstein. 1995b. A proposal for measuring scholarly productivity of accounting faculty. *Issues in Accounting Education* 10(2): 269-306.
- Hasselback, J.R., A. Reinstein, and M. Abdolmohammadi. 2012. Benchmarking the research productivity of accounting doctorates. *Issues in Accounting Education* 27(4): 943–978.
- Hasselback, J.R., A. Reinstein, and E.S. Schwan. 2000. Benchmarks for evaluating the research productivity of accounting faculty. *Journal of Accounting Education* 18: 79-97.
- Hasselback, J.R., A. Reinstein, and E.S. Schwan. 2003. Prolific authors of accounting literature. *Advances in Accounting* 20: 95–125.
- Huber, W.D. 2015. The research-publication complex and the construct shift in accounting research. *International Journal of Critical Accounting* 7(1): 1-48.

- Huber, W.D. 2016a. Deep impact: Impact Factors and accounting research. *International Journal of Critical Accounting* 8(1): 56-67.
- Huber, W.D. 2016b. Accounting research productivity: More heat than light?, *Journal of Theoretical Accounting Research* 11(2): 28-62.
- Kerr, S., M.G. Simkin, and R. Mason. 2009. Publication profile of accounting faculty. *American Journal of Business Education* 2(8): 103-114.
- Khurana, I.K, and K.K. Raman. 2003. Are fundamentals priced in the bond market? *Contemporary Accounting Research* 20(3): 465-494.
- Merriam-Webster. 2016. *Benchmark*. <https://www.merriam-webster.com/dictionary/benchmark>
- Reinstein, A., and T.A. Calderon. 2006. Examining accounting departments' rankings of the quality of accounting journals. *Critical Perspectives on Accounting* 17: 457-490.
- Richardson, A.J. and J.J. Williams. 1990. Canadian academic accountants' productivity: A survey of 10 refereed publications, 1976-1989. *Contemporary Accounting Research* 7(1): 278-294.
- Stone, D. (2017). *The "New Statistics" and Nullifying the null: Twelve actions for improving accounting research quality and integrity*. Paper presented to American Accounting Association Forensic Accounting Research Conference, Orlando, Florida, February 24-25, 2017.
- Walker, K.B., G.M. Fleischman, and T. Stephenson. 2010. The incidence of documented standards for research in departments of accounting at US institutions. *Journal of Accounting Education* 28(2): 43-57.
- Walker, K.B., G.M. Fleischman, and T. Stephenson. 2013. Developing a written research productivity policy for a department of accounting: A case study. *Academy of Educational Leadership* 17(3): 117-139.

Appendix A: Letter of Invitation

February 1, 2017

Dear Dean/Chairperson/Department Head:

A colleague and I are conducting a study of how accounting department administrators use published accounting faculty research productivity benchmarks to evaluate accounting faculty research productivity.

The study consists of an online survey which will take about 5-10 minutes to complete, depending on your responses.

We invite you to complete the survey which is available at <http://www.surveymonkey.com/r/CPKH5QG>

The survey will be available for four weeks beginning today.

Responses are anonymous, and you will have the option to request the results by sending me your email address.

Thank you for your valuable contribution to this study.

Sincerely,

Appendix 2: Survey

Welcome.

Welcome to the survey of benchmarks of accounting research productivity. Thank you for your participation. Your feedback is important.

The purpose of this survey is to determine the extent to which published benchmarks* of accounting research productivity are used in hiring, promotion, tenure, and merit pay decisions of accounting faculty.

There is no reward for completing the survey. There is no penalty for not completing the survey. It is strictly voluntary.

By completing the survey you understand and agree that there is neither reward nor penalty. The survey will take 5-10 minutes to complete, depending on your responses.

All responses are strictly anonymous and confidential.

*Here the term "benchmarks" is used to describe any measure of research productivity metrics or research productivity output

Background:

1. Are you a head, director, chair, dean, or other administrator of an accounting program, department, or school??
 - Yes
 - No
2. Is your university located in the U.S.?
 - Yes
 - No
3. What is the highest degree in accounting awarded by your university?
 - Bachelors
 - Masters (MA, MS, MBA, MAcc)
 - Doctorate (Ph.D., DBA)
4. Is your college/university accredited by the AACSB?
 - Yes
 - No
5. Is your accounting program accredited by the AACSB?
 - Yes
 - No
6. What is your position in your college/university?
 - Director/Head of Accounting Program
 - Chair of Accounting Department
 - Dean of Business/Accounting School
 - Other (please specify)
7. How long have you held your present position?
 - Less than 1 year.
 - 1-5 years.
 - 6-10 years.
 - 11-20 years.
 - More than 20 years
8. Is your college/university considered public or private?
 - Public
 - Private
9. Is your college/university considered sectarian/religious?
 - Yes
 - No

10. Is your college/university considered a research institution?

- Yes
- No

Benchmarks:

11. Are you familiar with the accounting research productivity benchmarks published by Hasselback and Reinstein (Assessing accounting doctoral programs by their graduates' research productivity, *Advances in Accounting*, 1995)?

- Yes
- No

12. Do you use the accounting research productivity benchmarks published by Hasselback and Reinstein (Assessing accounting doctoral programs by their graduates' research productivity, *Advances in Accounting*, 1995) to evaluate the accounting research productivity of accounting faculty?

- Yes
- No

13. How do you use the accounting research productivity benchmarks published Hasselback and Reinstein (Assessing accounting doctoral programs by their graduates' research productivity, *Advances in Accounting*, 1995) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)

- Tenure decisions
- Promotion decisions
- Merit pay decisions
- Hiring decisions
- Other (please specify)

14. Are you familiar with the accounting research productivity benchmarks published by Hasselback and Reinstein (A proposal for measuring scholarly productivity of accounting faculty, *Issues in Accounting Education*, 1995)?

- Yes
- No

15. Do you use the accounting research productivity benchmarks published by Hasselback and Reinstein (A proposal for measuring scholarly productivity of accounting faculty, *Issues in Accounting Education*, 1995) to evaluate the accounting research productivity of accounting faculty?

- Yes
- No

16. How do you use the accounting research productivity benchmarks published by Hasselback and Reinstein (A proposal for measuring scholarly productivity of accounting faculty, *Issues in Accounting Education*, 1995) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)

- Tenure decisions
- Promotion decisions
- Merit pay decisions
- Hiring decisions
- Other (please specify)

17. Are you familiar with the accounting research productivity benchmarks published by Hasselback, Reinstein, and Schwan (Benchmarks for evaluating the research productivity of accounting faculty, *Journal of Accounting Education*, 2000)?

- Yes
- No

18. Do you use the accounting research productivity benchmarks published by Hasselback, Reinstein, and Schwan (Benchmarks for evaluating the research productivity of accounting faculty, *Journal of Accounting Education*, 2000) to evaluate the accounting research productivity of accounting faculty?

- Yes
- No

19. How do you use the accounting research productivity benchmarks published by Hasselback, Reinstein, and Schwan (Benchmarks for evaluating the research productivity of accounting faculty, *Journal of*

- Accounting Education*, 2000) to evaluate the accounting research productivity of accounting faculty?
(Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)
20. Are you familiar with the accounting research productivity benchmarks published by Christensen, Finger, & Latham (New accounting scholars' publications in accounting and nonaccounting journals, *Issues in Accounting Education*, 2002)?
- Yes
 - No
21. Do you use the accounting research productivity benchmarks published by Christensen, Finger, & Latham (New accounting scholars' publications in accounting and nonaccounting journals, *Issues in Accounting Education*, 2002) to evaluate the accounting research productivity of accounting faculty?
- Yes
 - No
22. How do you use the accounting research productivity benchmarks published by Christensen, Finger, & Latham (New accounting scholars' publications in accounting and nonaccounting journals, *Issues in Accounting Education*, 2002) to evaluate the accounting research productivity of accounting faculty?
(Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)
23. Are you familiar with the accounting research productivity benchmarks published by Fogarty (Sustained research productivity in accounting: A study of the senior cohort, *Global Perspectives on Accounting Education*, 2004)?
- Yes
 - No
24. Do you use the accounting research productivity benchmarks published by Fogarty (Sustained research productivity in accounting: A study of the senior cohort, *Global Perspectives on Accounting Education*, 2004) to evaluate the accounting research productivity of accounting faculty?
- Yes
 - No
25. How do you use the accounting research productivity benchmarks published by Fogarty (Sustained research productivity in accounting: A study of the senior cohort, *Global Perspectives on Accounting Education*, 2004) to evaluate the accounting research productivity of accounting faculty?
(Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)
26. Are you familiar with the accounting research productivity benchmarks published by Everett, Klamm, & Stoltzfus (Developing benchmarks for evaluating publication records at doctoral programs in accounting, *Journal of Accounting Education*, 2004)?
- Yes
 - No
27. Do you use the accounting research productivity benchmarks published by Everett, Klamm, & Stoltzfus (Developing benchmarks for evaluating publication records at doctoral programs in accounting, *Journal of Accounting Education*, 2004) to evaluate the accounting research productivity of accounting faculty?

- Yes
 - No
28. How do you use the accounting research productivity benchmarks published by Everett, Klamm, & Stoltzfus (Developing benchmarks for evaluating publication records at doctoral programs in accounting, *Journal of Accounting Education*, 2004) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)
29. Are you familiar with the accounting research productivity benchmarks published by Glover, Prawitt, and Wood (Publication records of faculty promoted at the top 75 accounting research programs, *Issues in Accounting Education*, 2006)?
- Yes
 - No
30. Do you use the accounting research productivity benchmarks published by Glover, Prawitt, and Wood (Publication records of faculty promoted at the top 75 accounting research programs, *Issues in Accounting Education*, 2006) to evaluate the accounting research productivity of accounting faculty?
- Yes
 - No
31. How do you use the accounting research productivity benchmarks published Glover, Prawitt, and Wood (Publication records of faculty promoted at the top 75 accounting research programs, *Issues in Accounting Education*, 2006) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)
32. Are you familiar with the accounting research productivity benchmarks published by Reinstein and Calderon (Examining accounting departments' rankings of the quality of accounting journals, *Critical Perspectives on Accounting*, 2006)?
- Yes
 - No
33. Do you use the accounting research productivity benchmarks published by Reinstein and Calderon (Examining accounting departments' rankings of the quality of accounting journals, *Critical Perspectives on Accounting*, 2006) to evaluate the accounting research productivity of accounting faculty?
- Yes
 - No
34. How do you use the accounting research productivity benchmarks published by Reinstein and Calderon (Examining accounting departments' rankings of the quality of accounting journals, *Critical Perspectives on Accounting*, 2006) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)
35. Are you familiar with the accounting research productivity benchmarks published by Kerr, Simkin, and Mason (Publication profile of accounting faculty, *American Journal of Business Education*, 2009)?
- Yes

- No
36. Do you use the accounting research productivity benchmarks published by Kerr, Simkin, and Mason (Publication profile of accounting faculty, *American Journal of Business Education*, 2009) to evaluate the accounting research productivity of accounting faculty?
- Yes
 - No
37. How do you use the accounting research productivity benchmarks published by Kerr, Simkin, and Mason (Publication profile of accounting faculty, *American Journal of Business Education*, 2009) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)
38. Are you familiar with the accounting research productivity benchmarks published by Walker, Fleischman, & Stephenson (The incidence of documented standards for research in departments of accounting at US institutions, *Journal of Accounting Education*, 2010)?
- Yes
 - No
39. Do you use the accounting research productivity benchmarks published by Walker, Fleischman, & Stephenson (The incidence of documented standards for research in departments of accounting at US institutions, *Journal of Accounting Education*, 2010) to evaluate the accounting research productivity of accounting faculty?
- Yes
 - No
40. How do you use the accounting research productivity benchmarks published by Walker, Fleischman, & Stephenson (The incidence of documented standards for research in departments of accounting at US institutions, *Journal of Accounting Education*, 2010) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)
41. Are you familiar with the accounting research productivity benchmarks published by Hasselback, Reinstein, and Abdolmohammadi (Benchmarking the research productivity of accounting doctorates, *Issues in Accounting Education*, 2012)?
- Yes
 - No
42. Do you use the accounting research productivity benchmarks published by Hasselback, Reinstein, and Abdolmohammadi (Benchmarking the research productivity of accounting doctorates, *Issues in Accounting Education*, 2012) to evaluate the accounting research productivity of accounting faculty?
- Yes
 - No
43. How do you use the accounting research productivity benchmarks published by Hasselback, Reinstein, and Abdolmohammadi (Benchmarking the research productivity of accounting doctorates, *Issues in Accounting Education*, 2012) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)

44. Are you familiar with the accounting research productivity benchmarks published by Walker, Fleishman, and Stephenson (Developing a written research productivity policy for a department of accounting: A case study, *Academy of Educational Leadership*, 2013)?
- Yes
 - No
45. Do you use the accounting research productivity benchmarks published Walker, Fleishman, and Stephenson (Developing a written research productivity policy for a department of accounting: A case study, *Academy of Educational Leadership*, 2013) to evaluate the accounting research productivity of accounting faculty?
- Yes
 - No
46. How do you use the accounting research productivity benchmarks published Walker, Fleishman, and Stephenson (Developing a written research productivity policy for a department of accounting: A case study, *Academy of Educational Leadership*, 2013) to evaluate the accounting research productivity of accounting faculty? (Check all that apply.)
- Tenure decisions
 - Promotion decisions
 - Merit pay decisions
 - Hiring decisions
 - Other (please specify)

You have completed the survey of benchmarks of accounting research productivity. Thank you for your valuable participation. You may exit now. If you would like to receive the results, please email the researcher with your email address.

TABLE 1.
Data and descriptive statistics

Panel A. Background

Category	#(%)
Highest Degree in accounting:	
• Bachelors	47(19.50%)
• Masters	111(46.06%)
• Doctorate	83(34.44%)
Total	241(100%)
AACSB accredited?	
• Yes	191(79.25%)
• No	50(20.75%)
Total	241(100%)
Accounting AACSB accredited?	
• Yes	92(38.17%)
• No	149(61.83%)
Total	241(100%)
Public or private:	
• Public	146(60.58%)
• Private	95(39.42%)
Total	241(100%)
Sectarian or religious:	
• Yes	55(22.82%)
• No	186(77.18%)
Total	241(100%)
Research institution:	
• Yes	112(46.47%)
• No	129(53.53%)
Total	241(100%)

Panel B. Position and Years in Position

Category	#(%)
Position:	
• Dean of Business/Accounting School	46(19.09%)
• Chair of Accounting Department	129(57.68%)
• Director/Head of Accounting Program	46(19.09%)
• Other	23(9.54%)
Total	241(100%)
Years in Position:	
• Less than 1 year	23(9.54%)
• 1-5 years	127(52.70%)
• 6-10 years	61(25.31%)
• 11-20 years	24(9.96%)
• More than 20 years	6(2.49%)
Total	241(100%)

Panel C. Benchmark by Familiarity and Use

Published Benchmark	#(%) of total 241 familiar with published benchmarks	#(%) of total 241 not familiar with published benchmarks	#(%) of total 241 who actually use published benchmarks	#(%) total 241 who are familiar with published benchmarks who actually use published benchmarks
Hasselback and Reinstein (1995a)	116(48.13%)	125(51.87%)	8(3.32%)	8(6.90%)
Hasselback and Reinstein (1995b)	108(44.81%)	123(55.19%)	7(2.90%)	7(6.48%)
Hasselback, Reinstein and Schwan (2000)	100(41.49%)	141(58.51%)	4(1.66%)	4(4.00%)
Christensen, Finger and Latham (2002)	83(34.44%)	158(65.56%)	0(0%)	0(0%)
Fogarty (2004)	71(29.46%)	170(70.54%)	2(0.83%)	2(2.82%)
Everett, Klamm and Stoltzfus (2004)	71(29.46%)	170(70.54%)	1(0.41%)	1(1.41%)
Glover, Prawitt and Wood (2006)	106(43.98%)	135(56.02%)	18(7.47%)	18(16.98%)
Reinstein and Calderon (2006)	78(32.37%)	153(67.63%)	5(2.07%)	5(6.41%)
Kerr, Simkin and Mason (2009)	63(26.14%)	178(73.86%)	3(1.24%)	2(3.17%)
Walker, Fleischman and Stephenson (2010)	68(28.22%)	173(71.78%)	1(0.41%)	1(1.47%)
Hasselback, Reinstein and Abdolmohammadi (2012)	88(36.51%)	153(63.49%)	5(2.07%)	5(5.68%)
Walker, Fleishman and Stephenson (2013)	64(26.56%)	177(73.44%)	2(0.83%)	2(3.13%)

TABLE 2.
Chi-square analysis

Panel A. Demographic analysis

Demographic Variable	# and p-value who are familiar with Published Benchmarks (out of 241)	# and p-value who use Published Benchmarks (out of 241)
Position: $\chi^2(3, N = 241)$	19.46, p = .000***	5.15, p = .161
• Director/Head of Accounting Program	26(78.8%)	8(24.2%)
• Chair of Accounting Department	92(66.2%)	17(12.2%)
• Dean of Business/Accounting School	21(45.7%)	7(15.2)
• Other	7(30.4%)	1(4.3%)
Total	146(60.6%)	33(13.7%)
Years in present position: $\chi^2(4, N = 241)$	4.82, p = .306	7.5, p = .101
• Less than 1 year.	10(43.5%)	2(8.7%)
• 1-5 years.	76(59.8%)	23(18.1%)
• 6-10 years.	40(65.6%)	4(8.6%)
• 11-20 years.	17(70.8%)	2(8.3%)
• More than 20 years	3(50.0%)	2 33.3%)
Total	146(60.6%)	33(13.7%)
Highest degree of institution: $\chi^2(2, N = 241)$	48.51, p = .000***	9.40, p = .009***
• Bachelors	15(31.9%)	0(0.0%)
• Masters	57(51.4%)	18(16.2%)
• Doctorate	74(89.2%)	15(18.1%)
Total	146(60.6%)	33(13.7%)
AACSB accredited: $\chi^2(1, N = 241)$	9.34, p = .002***	10.50, p = .001***
• Yes	67(72.8%)	21(22.8%)
• No	79(53.0%)	12(8.1%)
Total	146(60.6%)	33(13.7%)
Public or private: $\chi^2(1, N = 241)$	3.12(p = .077)	5.31, p = .021***
• Public	95(65.1%)	26(17.8%)
• Private	51(53.1%)	7(7.4%)
Total	146(60.6%)	33(13.7%)
Sectarian or religious: $\chi^2(1, N = 241)$	2.79, p = .095	4.09, p = .043*
• Yes	28(50.9%)	3(5.5%)
• No	118(63.4%)	30(16.1%)
Total	146(60.6%)	33(13.7%)
Research institution: $\chi^2(1, N = 241)$	47.76, p = .000***	13.18, p = .000***
• Yes	94(63.9)	25(22.3%)
• No	52(40.3)	8(6.2%)
Total	146(60.69)	33(13.7%)

Panel B. Analysis by familiarity and use

Published Benchmark	Familiar with (out of 241)	Sig	Actually Use (out of 241)	Sig
Hasselback and Reinstein (1995a)	116(48.1%)	.562	8(3.3%)	.000***
Hasselback and Reinstein (1995b)	108(44.8%)	.107	7(2.9%)	.000***
Hasselback, Reinstein and Schwan (2000)	100(41.5%)	.008***	4(1.7%)	.000***
Christensen, Finger and Latham (2002)	83(34.4%)	.000***	0.0(0.0%)	.000***
Fogarty (2004)	71(29.5%)	.000***	2.8(1.2%)	.000***
Everett, Klamm and Stoltzfus (2004)	71(29.5%)	.000***	1(.4%)	.000***
Glover, Prawitt and Wood (2006)	106(44.0%)	.062	18(7.5%)	.000***
Reinstein and Calderon (2006)	78(32.4%)	.000***	5(2.1%)	.000***
Kerr, Simkin and Mason (2009)	63(26.1%)	.000***	3(1.24%)	.000***
Walker, Fleischman and Stephenson (2010)	68(28.2%)	.000***	1(.4%)	.000***
Hasselback, Reinstein, Abdolmohammadi (2012)	88(38.5%)	.000***	5(2.1%)	.000***
Walker, Fleishman and Stephenson (2013)	64(26.6%)	.000***	2(.8%)	.000***

Panel C. Analysis by decisions

Decision	Tenure (out of 241)	Promotion (out of 241)	Merit Pay (out of 241)	Hiring (out of 241)
Hasselback and Reinstein (1995a)	8(3.3%) $p = .000***$	8(3.3%) $p = .000***$	6(2.5%) $p = .000***$	1(0.4%) $p = .000***$
Hasselback and Reinstein (1995b)	7(2.9%) $p = .000***$	7(2.9%) $p = .000***$	4(1.4%) $p = .000***$	1(0.4%) $p = .000***$
Hasselback, Reinstein and Schwan (2000)	3(1.2%) $p = .000***$	3(1.2%) $p = .000$	1(0.4%) $p = .000***$	0(0.0%) ^a
Christensen, Finger and Latham (2002)	0(0.0%) ^a	0(0.0%) ^a	0(0.0%) ^a	0(0.0%) ^a
Fogarty (2004)	2(0.8%) $p = .000***$	2(0.8%) $p = .000***$	1(0.4%) $p = .000***$	2(0.8%) $p = .000***$
Everett, Klamm and Stoltzfus (2004)	1(0.4%) $p = .000***$	1(0.4%) $p = .000***$	0(0.0%) ^a	0(0.0%) ^a
Glover, Prawitt and Wood (2006)	13(5.4%) $p = .000***$	13(5.4%) $p = .000***$	4(1.7%) $p = .000***$	5(2.1%) $p = .000***$
Reinstein and Calderon (2006)	5(2.1%) $p = .000***$	$p = .000***$	2(.08%) $p = .000***$	2(0.8%) $p = .000***$
Kerr, Simkin and Mason (2009)	1(0.4%) $p = .000***$	1(0.4%) $p = .000***$	2(0.8%) $p = .000***$	1(0.4%) $p = .000***$
Walker, Fleischman and Stephenson (2010)	0(0.0%) ^a	0(0.0%) ^a	1(0.4%) $p = .000***$	0(0.0%) ^a
Hasselback, Reinstein , Abdolmohammadi (2012)	8(3.3%) $p = .000***$	7(2.9%) $p = .000***$	5(2.1%) $p = .000***$	1(0.4%) $p = .000***$
Walker, Fleishman and Stephenson (2013)	1(0.4%) $p = .000***$	1(.04%) $p = .000***$	1(0.4%) $p = .000***$	0(0.0%) ^a

^aThis variable is constant (0). χ^2 test could not be performed.

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