THE ACCOUNTING EDUCATORS' JOURNAL
Volume XXXIII
2023
pp. 189 - 217

Causal Effect of Business Accreditation on the CPA Exam Success Rate

Pamela Baker
Texas Woman's University

Robert Maurer
Texas Woman's University, Professor Emeritus

Zhen Li
Texas Woman's University

Lin Zou
Texas Woman's University

Mark Tengesdal
Texas Woman's University

Abstract

We examine the relationship between AACSB post-secondary business program accreditation and program success rate on the CPA Exam. We use panel data and a difference-in-differences quasi-experimental design to identify the causal effect of AACSB accreditation on the CPA exam success rate. In programs with intermediate to large numbers of exam candidates, we find no evidence that accreditation *per se* is causally related to the pass rate although a positive correlation between them is supported. In programs with small numbers of exam candidates, our results show the existence of causal effect of accreditation on the pass rate.

Keywords: Business accreditation, CPA exam, Difference-in-differences, Causal inference

Introduction

In higher education, accreditation is a performance measure that programs use to signal quality and that prospective students and employers use to make educational investment and hiring decisions (Roller et al., 2003; Kundu, 2020). The Council for Higher Education Accreditation recognizes three organizations that offer accreditation to postsecondary business programs: the Association to Advance Collegiate Schools of Business (AACSB), the Accreditation Council for Business Schools and Programs (ACBSP), and the International Assembly for Collegiate Business Education (IACBE). These accrediting bodies differ primarily with respect to cost and the rigor and rigidity of their accreditation standards (Brink & Smith, 2012).

Similarly, in many professional fields, a licensure exam regulates entry to the field and a program's success rate on the exam can be interpreted as a signal that reveals information about the program's absolute and/or relative level of quality. In the accounting profession, the success rate on the Uniform Certified Public Accountants (CPA) Exam is frequently interpreted in this way.

Prior research (outlined in the next section) has found a positive correlation between AACSB accreditation and program pass rates on the CPA exam. This is to be expected since both accreditation status and the pass rate convey complementary information about program quality. However, the causal direction of the relationship is an open question. One hypothesis consistent with positive correlation, for example, posits that high-pass rate programs pursue AACSB accreditation to signal their quality in a competitive market. An alternative hypothesis consistent with positive correlation posits that low-pass rate programs also pursue AACSB accreditation to enhance their quality reputation in the market; however, in this case, securing accreditation requires additional investments in faculty and educational resources which then improve exam pass rates. In both cases, the correlation between accreditation status and the exam pass rate is positive but the causal directions are reversed. In the former, a high-pass rate motivates the pursuit of accreditation but, in the latter, the pursuit of accreditation motivates an improvement in the pass rate.

We try to distinguish between these cases. Specifically, we test the hypothesis that the pursuit of AACSB accreditation influences the program pass rate on the CPA exam. This hypothesis was motivated by interviews we conducted with business programs that had recently received accreditation from AACSB. In the course of these interviews, we found that programs frequently needed to increase the size of their accounting faculty complement to meet AACSB expectations and that the criteria used to qualify faculty were, in general, more rigorous than the criteria used by either ACBSP or IACBE.

The causal direction of the relationship is an important policy consideration. The fact that program quality is difficult to observe and measure implies that programs must use market signals, such as accreditation, to communicate quality to stakeholders; the market structure that ensues has the structure of an asymmetric information signaling game and can result in complex and counterintuitive incentives (Connelly et al., 2011). For example, Boleslavsky and Cotton (2015) model the generation of information signals in the interaction between grading standards and program investments in quality. In their model, competitive market conditions supply programs with incentives to engage in "strategic grading" in which programs issue less than perfectly informative performance assessments (i.e., grade inflation) to attract students. Prospective employers, who need a more informative signal about the quality of program graduates, rely instead on alternative signals (such as accreditation) that, in turn, supply programs with incentives to invest in observable attributes that are correlated with quality. Thus, counterintuitively, lowering grading standards motivates investments in reputational signals such as accreditation.

We use panel data and a difference-in-differences quasi-experimental design to identify the causal direction of the relationship between AACSB accreditation and program pass rates on the CPA exam. The data sample consists of pass rates for first-time candidates from programs that received initial accreditation from AACSB (the treatment group) or ACBSP (the control group) between 1990 and 2009. In order to reduce bias introduced by the small-cohort rate-cluster effect, we consider programs with small number of candidates and those with intermediate to large number of candidates separately. It is consistent with Jantzen and Pendleton's (1994) finding that the prospects for AACSB accreditation was related to the size of program. For mid-/large-size programs, although we find a positive association between the pass rate and membership in the AACSB group, we find no evidence that this pass rate effect is causally related to AACSB accreditation. However, for small-size programs, we find that AACSB accreditation *per se* was causally related to changes in the exam pass rate relative to ACBSP accreditation.

In the next section, we review the existing literature on the relationship between program accreditation and the CPA exam. In the section that follows, we describe the data used in our study, the difference-in-differences research design, and the structure of the causal inference problem. In particular, we focus on a number of important econometric issues and how we addressed them. In the following section, we present our empirical findings. And, in the final section, we present our conclusions which include a discussion of the study's limitations and suggestions for future research.

Literature Review

Existing research on the relationship between business school accreditation and CPA exam pass rates is somewhat mixed. In an early study, Marts et al. (1988) compared CPA exam pass rates between programs with AACSB accounting accreditation, programs with AACSB business (but not accounting) accreditation, and programs without AACSB accreditation. They found that programs with AACSB accounting accreditation had higher pass rates than programs without AACSB accreditation but found no significant difference in the pass rate between AACSB programs with accounting accreditation and AASCB programs with only business accreditation. In a similar study, conducted two and a half decades later with additional data, Bunker et al. (2014) also found that programs with AACSB accounting accreditation had higher pass rates, and higher average exam scores, than programs with only AACSB business accreditation. More recently, Nagle et al. (2018) confirmed the Bunker et al.'s (2014) findings and further reported that business-only-accredited programs also outperformed the unaccredited programs.

Barilla et al. (2008) used logit regression to repeat the Marts et al. (1988) study finding that AACSB accounting accreditation increased the odds of passing all CPA exam sections by first-time candidates. They also studied the performance of programs with AACSB business-only accreditation, ACBSP accreditation, and IACBE accreditation finding that none of the three increased the odds of success on the exam.

Morgan et al. (2008) compared CPA exam success rates between AACSB programs with business accreditation and non-AACSB programs; while Bergin et al. (2011) extended the analysis to a comparison between AACSB programs, ACBSP programs, and programs with neither accreditation. Both studies found significantly higher success rates associated with AACSB programs compared to ACBSP programs and programs with neither accreditation. Morgan et al. (2012) used mean CPA exam scores, rather than pass rates, as the dependent variable and extended the analysis to a comparison between AACSB programs, ACBSP programs, IACBE programs, and programs with no accreditation. Similar to studies that use pass rates as the dependent variable, they found that first-time candidates from AACSB programs have significantly higher mean scores than candidates from either ACBSP programs, IACBE programs, or programs with no accreditation. However, counterintuitively, they also found that candidates from ACBSP programs and IACBE programs tended to perform worse than candidates from programs with no accreditation.

In 1988, the American Institute of Certified Public Accountants (AICPA) adopted a rule increasing the number of educational credits required for CPA licensure to 150 hours (Fuller, 2008). This requirement, which went into effect at different times in different jurisdictions, generated a debate about whether the additional credit hours affected the success rate on the CPA exam (Gaynor et al., 2019). Several studies that examined this issue included program accreditation as a control variable in the analysis. Grant et al. (2001) found that three additional accounting credit hours had a relatively weak effect on the probability of success, but that AACSB accreditation had a relatively strong effect. On the other hand, Boone et al. (2006), whose results were otherwise consistent with the Grant et al. (2002) study, found only weak evidence for an AACSB accreditation effect. Since these two studies included controls for the influence of program characteristics, such as performance on standardized entrance exams, their results suggest that the association between AACSB accreditation and enhanced success on the exam may result, not from AACSB accreditation *per se*, but from the influence of unobserved variables, such as higher quality students or more faculty resources, that are correlated with AACSB accreditation.

Overall, the existing evidence supports a correlation between AACSB accreditation and enhanced success on the CPA exam (Albring & Elder, 2020). However, with the possible exception of the Grant et al. (2002) and the Boone et al. (2006) studies (which include controls for individual program characteristics), this evidence cannot be used to support causal inferences about the relationship between accreditation and observed differences in exam success rates. Morgan et al. (2012), for example, emphasize this limitation stating that their findings do not support causal inference because the research design does not use random assignment (to control for the influence of omitted variables).

Methods

Data Sample

The data used in our study was obtained from the National Association of State Boards of Accountancy (NASBA) for the years 1985-2014. Since the only outcome measure available for the entire period was the program pass rate on the exam and its subsections, we used the program pass rate as our dependent variable. However, over this period, the exam's administration, organization, and content changed a number of times. Since these changes have been discussed in detail elsewhere (Bergin et al., 2011; Bunker et al., 2014), we focus here primarily on the issues these changes presented for our analysis.

Between 1985 and 2003, the exam was administered in May and November using a paper-and-pencil format (NASBA, 2005). In 2004, NASBA and the AICPA changed the exam's administration to a computer-based format offered continuously for a two-month window each quarter. This change affected how the exam's results were reported. Prior to 2004, exam pass rates were reported separately for the May and November administrations. Beginning in 2004, and for subsequent years, pass rates were reported by year of administration. Prior to 2004, candidates also sat for all four exam sections at a single administration and were required to pass at least two sections to earn completion credits. With the change to a computer-based format in 2004, candidates were permitted to take, and to receive credits, one section at a time. Thus, pass rates reported prior to 2004 are not necessarily comparable to pass rates reported after 2004.

Similarly, the exam's content and organization evolved over the study period. During the 1985-1993 period, the exam was organized into four sections: Auditing, Law, Theory, and Practice. In 1994, a number of significant changes were made to the exam (NASBA, 1995). These included a substantial reduction in the exam's duration and reorganization of the Theory and Practice sections. After the reorganization, the sections were designated: Auditing (AUD); Business Law & Professional Responsibilities (LPR); Financial Accounting and Reporting -- Business Enterprises (FARE); and Accounting and Reporting -- Taxation, Managerial, and Governmental and Not-For-Profit Organizations (ARE). In 2004, when the exam's administration was changed to a computer-based format, simulation questions were introduced and the four sections were renamed (NASBA, 2005): Auditing and Attestation (AUD); Business Environment and Concepts (BEC, formerly LPR); Financial Accounting and Reporting (FAR, formerly FARE); and Regulation (REG, formerly ARE). In 2011, certain written communication tasks were dropped and simulation questions changed.

Pass Rate Construction

The program pass rates were constructed as follows. As noted above, pass rates prior to 2004 were reported separately for the May and November administrations but, after 2004, were reported on an annual basis. Pass rates were further differentiated by whether an individual was a first-time candidate or was repeating the exam subsequent to an earlier attempt and, also, by whether a candidate held an advanced degree beyond the bachelor's level. Since our study focused on performance conditional on type of accreditation, we elected to use only pass rates for first-time candidates. We excluded repeat candidates because this group: (1) includes candidates with exam experience at a previous administration, (2) generally are further removed in time from their educational programs than first-time candidates, and (3) are more likely to have received additional training or experience in accounting through their place of employment. Thus, like most previous studies, pass rates for first-time candidates were thought to be a better measure of the effect of program accreditation on student performance.

To construct the 1985-2003 pass rate series, we compiled pass rates for the May and November exam administrations reported in Table 12A (first-time candidates without an advanced degree) and Table 12B (first-time candidates with an advanced degree) from the annual *CPA Candidate Performance* reports for the programs in our study sample (NASBA, 1985-2004). For each program, we constructed the annual pass rate in a given year from the May and November pass rates using the number of candidates at each administration as weights. We then combined the weighted average annual pass rates for 1985-2003 with the annual pass rates by program for the 2004-2014 period supplied to us in electronic form by NASBA. Since a *CPA Candidate Performance* report for 1988 was not available, our study does not include data for that year.

Econometric Issues

Three econometric issues were associated with the pass rate data and periodic changes in the exam's administration, organization, and content. The first involved time trends in the pass rates; the second involved clustering in the observations; and the third involved measurement issues.

Time Trend Issues

Figure 1 presents time trends in the section pass rates for first-time candidates. Prior to the late 1990's, the average annual pass rate on all four sections ranged between 25% and 35%. Around the year 2000, average pass rates on the four sections began a sustained rise. Since 2012, the pass rate on the BEC section has been above 60% while the pass rates on the AUD, FAR, and REG sections have ranged between 45% and 55%.

These time trends created a problem for our difference-in-differences design because pass rate observations are assigned to the time index relative to a program's year of accreditation (e.g., 1-year prior to year of accreditation, 2-years prior, and etc.) rather than the calendar year in which the observation occurs. Since the difference-in-differences design assumes that both the control group and the treatment group are subject to the same systematic influences, referred to as the "parallel trends" assumption (Angrist & Pischke, 2009), and observations in the same relative year can differ with respect to the calendar year from which the observations are drawn, the time trends in Figure 1 can introduce variation that violates this assumption.

To control for these year-to-year trends, we rescaled the program pass rate in a given year relative to the mean pass rate for that section in that year. Thus, in a given year, we defined the relative pass rate as a program's observed pass rate for first-time candidates divided by the population mean pass rate for first-time candidates sitting for the exam in that year. This rescales the mean rate to 1.0 in all years.

Pass Rate Clustering Issues

The second issue was related to the use of program pass rates as the outcome measure. Figure 2 presents the distribution of program pass rates by number of candidates (Panel A) and by average candidate score (Panel B) for the 2014 AUD section of the exam. Panel A shows that program pass rates cluster at certain intervals when the number of candidates from a program is small. For example, when only one student from a program sits for the exam, the possible pass rates are the binary extremes 0% and 100%; when two students from a program sit for the exam, the possible pass rates are 0%, 50%, and 100%; when three students sit for the exam, the possible pass rates are 0%, 33%, 67%, and 100%; and so on. Panel B shows the extent of the clustering problem. In Panel B, note that the pass rate interval endpoint "attractors" associated with small candidate cohorts in Panel A have abnormally wide average score ranges.

In statistical terms, the program pass rate behaves like a censored variable in the sense that a wide range of scores map to a limited number of pass rates. Since pass rates map to the endpoints of intervals with length $1/n$, where n is the cohort size, the interval between pass rates has an upward bias inversely related to cohort size. Thus, including small cohorts in the analysis has the potential to introduce a significant source of bias into the statistical estimates.

To control for this clustering problem, we filtered the programs in our sample by the number of candidates sitting for the exam. The dashed lines in Panel A of Figure 2 suggest that this effect is reduced to the level of background noise at a threshold of 15 to 20 candidates. The criteria we used to filter the data is described in greater detail in the *Design and Implementation* section below.

Measurement Issues

The periodic revisions to the exam also were the source of measurement issues. The basic problem can be illustrated with a simple example. Suppose the exam contains two sections with one question in Section 1 and two questions in

Section 2. Further, suppose Program H (high performing) and Program L (low performing) have an equal number of candidates sitting for the exam and all candidates from a given program provide the same right or wrong answer to each of the three questions. When the exam is administered, Program H students answer all three exam questions correctly resulting in a 100% pass rate on both sections; Program L students, however, answer only a single question in Section 2 correctly resulting in a 0% pass rate on Section 1 and a 50% pass rate on Section 2. In a subsequent period, the exam is reorganized and the Section 2 question that Program L students answered incorrectly is moved to Section 1. The reorganization does not affect Program H's pass rates, which remain 100% on both sections, but Program L's pass rate on Section 2 improves from 50% to 100% while remaining at 0% on Section 1. Thus, even though there was no substantive change in performance for either program, the effect of the reorganization was an improvement in Program L's Section 2 pass rate.

Some of the trend changes in Figure 1 could reflect a composition effect of this type. In particular, the apparent trend discontinuity between the years 2003 and 2004 could reflect the extensive changes that were made to the exam in 2003. Similarly, the incremental evolution over time in the exam's administration, organization, and content (Audit → AUD → AUD; Law → LPR → BEC; Theory → FARE → FAR; Practice → ARE → REG) raises questions about what the program pass rates measure and whether pass rates from one period can be compared to pass rates from more distant periods.

As a performance metric, the program pass rate can be interpreted in at least two ways. When the exam's administration, organization, and content remain consistent from one period to the next, the pass rate measures program performance relative to that exam regime. However, when the exam regime changes over time, the pass rate measures program performance relative to the accounting standards and practices that the exam is designed to reflect. In other words, as the accounting industry evolves, adopting new standards and practices or revising and discarding old ones, the exam regime evolves to reflect these changes. Similarly, educational programs that prepare students for the exam will respond to changes in the regime with changes in the program's emphasis and curriculum. Consequently, program pass rates reflect the outcome of a dynamic adaptive process in which programs evolve in response to an evolving exam regime.

For instance, in the example above, Program L's performance improvement should be interpreted using the latter interpretation. Under that interpretation, Program L's improvement reflects a change in how the industry organizes the exam's content rather than a change in program performance. Since the reorganization affects the pass rates of other programs, it will also affect the mean program pass rate in the program population. Thus, measuring performance relative to the mean should, to some extent, compensate for this composition effect.

We adjust for this dynamic process in two respects. First, we rescale the program pass rate to measure performance on a given exam section relative to the mean population pass rate on that section in the same year. This adjusts the pass rate for systematic trends in the average pass rate due to changes in the exam regime. Second, we use the difference-in-differences quasi-experimental design to test if the pass rate has been changed before/after the accreditation year. Our design assigns the pass rate observations into three sub-periods based on the accreditation year: pre-accreditation, phase-in, and post-accreditation. The effect of accreditation in this design depends on the before/after accreditation distinction.

We also controlled for composition effects using a composite pass rate R_i^* which we defined as the equally-weighted 2nd-order mean of a program's pass rates on the four exam sections,

$$R_i^* = \sqrt{\sum_m \frac{1}{4} R_{im}^2} \tag{1}$$

where R_{im} denotes the pass rate for the i-th program and section $m \in \{\mathrm{AUD}, \mathrm{BEC}, \mathrm{FAR}, \mathrm{REG}\}$. This construction has a convenient geometric interpretation. If the section pass rates represent orthogonal outcomes in a four-dimensional vector space, which we think is reasonable since the exam treats the sections as independent skill sets, then R_i^* is the Euclidean length of a vector constructed from the individual section pass rates normalized to unit length. Figure 3 presents a graphical illustration using the pass rates on three sections (since the graphic is limited to three spatial dimensions).

This definition has the convenient property that R_i^* is homogeneous of degree one. Thus, scaling the section pass rates by some common factor scales the composite pass rate by the same factor. In addition, R_i^* belongs to the Constant Elasticity of Substitution class of functions, which is used to model firm production functions and consumer utility functions in economic applications, and implies that changes in section pass rates, due to changes in the exam, generate similar percentage changes in the composite rate.

Difference-In-Differences Estimation

The identification of causal relationships is an important econometric problem and has received a great deal of attention in recent years (Angrist & Pischke, 2009; Angrist & Pischke, 2010; Panhans & Singleton, 2017). In this section, we explain the causal identification strategy used by the difference-in-differences method.

To illustrate the nature of the problem, let R_{1i} denote the pass rate on the CPA exam for program *i* when it has AACSB accreditation and R_{0i} denote the pass rate in the alternative state. Note that R_{1i} and R_{0i} denote *potential* outcomes that depend on the program's accreditation choice. The *observed* pass rate R_i, expressed as a function of the potential outcomes and the program's accreditation choice, is given by

$$R_i = R_{0i} + (R_{1i} - R_{0i})d_i \tag{2}$$

where d_i denotes the program's choice and takes the value 1 when program *i* chooses AACSB accreditation and the value 0 otherwise.

In this setting, the causal relationship between AACSB accreditation and the pass rate is the difference in the potential outcomes R_{1i} - R_{0i}. To measure this difference, collect a sample of programs and estimate the difference in potential outcomes by the difference in the average observed pass rate for each of the choice groups,

$$E[R_i|d_i = 1] - E[R_i|d_i = 0] = E[R_{1i}|d_i = 1] - E[R_{0i}|d_i = 0]. \tag{3}$$

Expanding the right-hand side of the equation, however, shows that this estimation strategy is subject to selection bias when the treatment (AACSB accreditation) is non-random,

$$\underbrace{E[R_i|d_i = 1] - E[R_i|d_i = 0]}_{\text{observed difference in the average pass rates}} = \underbrace{E[R_{1i} - R_{0i}|d_i = 1]}_{\text{average effect of accreditation}} + \underbrace{\{E[R_{0i}|d_i = 1] - E[R_{0i}|d_i = 0]\}}_{\text{selection bias}}. \tag{4}$$

The second term on the right-hand side is the difference in the average untreated outcome between the two choice groups. This term leads to selection bias when systematic differences between the groups result in differences in the average pass rates.

Suppose, for example, that programs with superior resources choose AACSB accreditation to signal their quality and, thus, attract more high performing students relative to non-AACSB programs. In that case, the untreated outcome R_{0i} depends on a program characteristic (superior resources) that is correlated with the treatment (AACSB accreditation) but is not causally related to the treatment. Since an observed difference in pass rates between the choice groups depends on both an accreditation effect and a selection effect, one must control for the selection effect to identify and estimate the causal effect of accreditation.

In biostatistics, the "gold standard" design used to identify causal effects relies on random assignment of the treatment to eliminate selection bias which, in the previous example, would sever the link between resources and accreditation. In social sciences, however, random assignment is rarely possible and researchers must rely on alternative methods. One such method is the difference-in-differences identification strategy which uses changes in outcome levels, rather than the levels themselves, to estimate causal effects. When both the treatment group and the control group are exposed to a common set of influences then, in the absence of treatment, both groups experience similar changes in the outcome level, even though the two groups might begin at different levels initially. When the treatment group trend diverges from this common trend, then this is evidence for a treatment effect.

To illustrate, suppose we have pass rate data for both choice groups at two time points: (1) a pre-treatment time point prior to the AACSB program receiving accreditation, and (2) a post-treatment time point after the AACSB program received accreditation. Using superscripts to indicate time points, the difference-in-differences estimator δ can be written as

$$\delta = E\left[R_i^{\text{post}} - R_i^{\text{pre}} | d_i = 1\right] - E\left[R_i^{\text{post}} - R_i^{\text{pre}} | d_i = 0\right]. \tag{5}$$

This estimator is the difference between the pre to post changes in the pass rate for the two choice groups. Expanding the right-hand side, as above, shows that the bias term has a somewhat different interpretation,

$$\delta = \underbrace{E\left[\left(R_{1i}^{\text{post}} - R_{1i}^{\text{pre}}\right) - \left(R_{0i}^{\text{post}} - R_{0i}^{\text{pre}}\right) | d_i = 1\right]}_{\text{average effect of accreditation}} + \underbrace{\left\{E\left[R_{0i}^{\text{post}} - R_{0i}^{\text{pre}} | d_i = 1\right] - E\left[R_{0i}^{\text{post}} - R_{0i}^{\text{pre}} | d_i = 0\right]\right\}}_{\text{parallel trends term}}. \tag{6}$$

In this case, the second term is a common trend, or parallel trends, assumption. Note that, when both groups share a common trend in the absence of treatment, the parallel trends term vanishes and the difference-in-differences estimator provides an estimate of the causal effect of the treatment. Written in terms of levels, the difference-in-differences estimator takes the form

$$\delta = \underbrace{E\left[R_{1i}^{\text{post}} | d_i = 1\right]}_{\substack{\text{post-treatment pass rate} \\ \text{in AACSB group}}} - \underbrace{\left\{E\left[R_{0i}^{\text{post}} | d_i = 0\right] + \left\{E\left[R_{1i}^{\text{pre}} | d_i = 1\right] - E\left[R_{0i}^{\text{pre}} | d_i = 0\right]\right\}\right\}}_{\text{counterfactual pass rate in AACSB group}}. \tag{7}$$

In this form, the second term represents a *counterfactual* pass rate: the post-treatment time point pass rate in the AACSB group had it not experienced accreditation.

Control Group

In a study designed to detect the causal effect of business program accreditation on the CPA exam pass rate, a natural choice for the control group would be programs lacking business accreditation. In our study, we use AACSB programs as the treatment group but use ACBSP programs, rather than unaccredited programs, as the control group. We made this choice for three reasons.

First, ideally, the treatment group and control group should differ only in the application of the treatment. As noted in the *Introduction*, interviews we conducted with programs that had recently received AACSB accreditation suggested that substantial differences in resources devoted to accounting existed between AACSB programs on the one hand and ACBSP, IACBE, and unaccredited programs on the other. Given this difference, we felt the dividing line between treatment (an increase in resources devoted to accounting and more rigorous faculty standards) and no treatment was between AACSB programs and ACBSP/IACBE/no-accreditation programs.

Second, in a difference-in-differences design, the parallel trends assumption requires that the difference between the pass rates in the treatment group and the control group be a function of only factors related to the treatment. Or, stated differently, both groups should exhibit identical responses to environmental shocks in the absence of the treatment. The environment for a non-accredited program (or IACBE-accredited, which is essentially a very weak accreditation) would be quite different from the environment for an AACSB-accredited program. Therefore, they have different abilities and resources to respond to the uncertainty caused by physical and sociocultural changes. However, in order to prepare for either AACSB or ACBSP accreditation, programs are required to seek links with the local community and demonstrate the societal impact at a local, regional, national or international level (www.aacsb.com; www.acbsp.org). They are more likely to respond to environmental shocks in a similar way. We felt that the ACBSP group was more likely to satisfy this assumption.

Third, from a practical standpoint, most unaccredited programs lack sufficient candidates to pass the filter we use to control for the pass rate cluster issue described above. Similarly, IACBE programs have the same insufficient candidate problem as unaccredited programs. As a result, it left ACBSP programs as the only feasible control group.

In light of the above considerations, we chose ACBSP programs as the control group in this study. In the accounting field, success on the CPA exam is frequently used as one significant measure of program quality (Franklin & Myers, 2016). The quality of the business program varies with the contents of the accreditation standards and the levels of rigor and rigidity of its guideline. Thus, we expect that the differences between AACSB and ACBSP accreditations will influence the CPA exam performance. For example, AACSB has a more rigorous definition of qualified faculty, requires a higher percentage of qualified faculty, and has higher standards for scholarship as compared to ACBSP (Brink & Smith, 2012). These factors have been identified as important accounting faculty characteristics that have a significant impact on candidate's CPA exam performance (Bline et al., 2016).

Design and Implementation

Our study used panel data with a difference-in-differences quasi-experimental design. This design compares the pass rate on the CPA exam after accreditation in a treatment group (AACSB programs) with a counterfactual post-accreditation pass rate constructed using a control group (ACBSP programs).

In our design, post-accreditation observations were defined as the pass rates in the five years following the year in which the program received accreditation. The pre-accreditation observations were drawn from the ten years prior to, and including, the year in which the program received accreditation. Since AACSB accreditation often requires changes to the program that require a number of years to implement, the first five years of this ten-year period were designated as pre-accreditation observations while the latter five years, which end with the year of accreditation, were designated as an accreditation preparation and phase-in period. This structure is illustrated in Figure 4; Panel A presents the basic structure of our design and Panel B illustrates the role of the parallel trends assumption.

To address the pass rate clustering issue discussed above, we filtered the pass rate observations for data records with sufficient numbers of candidates. As described above, the unit of analysis consisted of fifteen years of pass rate observations divided into three five-year subperiods. Our data filter operated in two stages. In the first stage, we counted the number of years in each subperiod with candidate counts that exceeded a specified minimum. In the second stage, we specified a second threshold for the number of years in the subperiod with minimum candidate counts in the first stage. If all three subperiods satisfied the specified minimums, then the fifteen-year pass rate data record for that program was included in the data used in the analysis.

We implemented the difference-in-differences analysis using ordinary least squares regression (Angrist & Pischke, 2009). Let R_{ist} be the pass rate on the CPA exam for program *i* and accreditation status *s* in period *t* where $t \in \{-9, -8, ..., 0, ..., +5\}$ denotes the year relative to the year in which the program was granted accreditation. Let d_s be a state dummy and d_t a time dummy such that

$$d_s = \begin{cases} 0 \text{ for ACBSP} \\ 1 \text{ for AACSB} \end{cases} \tag{8}$$

and

$$d_t = \begin{cases} 0 \text{ for } t \in \{-9, -8, ..., -5\} \\ 1 \text{ for } t \in \{+1, +2, ..., +5\} \end{cases} \tag{9}$$

The pre-accreditation to post-accreditation difference-in-differences analysis, formulated using a linear regression model, is given by the equation

$$R_{ist} = \beta_0 + \beta_s d_s + \beta_t d_t + \delta d_s d_t + \varepsilon_{ist}. \tag{10}$$

This model says that the predicted pass rate on the exam is the sum of β_0, the expected pre-accreditation pass rate in the ACBSP control group, plus three effects: (1) β_s the pre-accreditation effect of program membership in the

AACSB group relative to membership in the ACBSP control group, (2) β_t the post-accreditation effect on the pass rate of membership in the ACBSP control group, and (3) δ the differential influence of AACSB accreditation on the post-accreditation change in the pass rate relative to ACBSP accreditation. Or, in econometric terms,

$$\beta_0 = \underbrace{E[R_{ist} | d_s = 0, d_t = 0]}_{\text{pre-accreditation pass rate in the ACBSP control group}} \tag{11}$$

$$\beta_s = \underbrace{E[R_{ist} | d_s = 1, d_t = 0] - E[R_{ist} | d_s = 0, d_t = 0]}_{\text{pre-accreditation effect of AACSB relative to ACBSP}} \tag{12}$$

$$\beta_t = \underbrace{E[R_{ist} | d_s = 0, d_t = 1] - E[R_{ist} | d_s = 0, d_t = 0]}_{\text{post-accreditation effect in the ACBSP control group}} \tag{13}$$

$$\delta = \underbrace{E[R_{ist} | d_s = 1, d_t = 1]}_{\text{AACSB post-accreditation pass rate}} - \underbrace{\{E[R_{ist} | d_s = 0, d_t = 1] + \{E[R_{ist} | d_s = 1, d_t = 0] - E[R_{ist} | d_s = 0, d_t = 0]\}\}}_{\text{counterfactual AACSB post-accreditation pass rate}} \tag{14}$$

The δ coefficient on the interaction term in the regression model is the difference-in-differences estimate constructed from the means of the pre-accreditation and post-accreditation periods.

Since the structure of our study included a phase-in period, in addition to the pre-accreditation and post-accreditation periods, we modified the regression model as follows,

$$R_{ist} = \beta_0 + \beta_s d_s + \beta_{t_2} d_{t_2} + \beta_{t_3} d_{t_3} + \delta_{t_2} d_s d_{t_2} + \delta_{t_3} d_s d_{t_3} + \varepsilon_{ist}. \tag{15}$$

where the subscripts t_2 and t_3 denote the phase-in period (subperiod 2) and the post-accreditation period (subperiod 3). The model does not include a term for the pre-accreditation period (subperiod 1) to avoid collinearity due to a design matrix with less than full column rank. In this model, the coefficient on the t_3 interaction term is equivalent to the δ coefficient in the basic model above (Eq. 10). The coefficient on the t_2 interaction term is the difference-in-differences estimate constructed from the means of the phase-in and post-accreditation periods.

Results

Data Characteristics

The data sample consisted of pass rate and candidate count observations from the 1985-2014 period for 504 programs that received initial accreditation from either AACSB (204 programs) or ACBSP (300 programs) between 1990 and 2009. One program received AACSB and ACBSP accreditation at different times during the study period and is included in both data sets. Figure 5 presents the distribution of accreditation type by year of accreditation. Although the ratio of AACSB to ACBSP programs in Figure 5 changes from year to year, no clear pattern or trend is evident in these changes.

The overall characteristics of the sample are presented in Table 1. Depending on the section, 191-193 out of 204 AACSB programs (93%-94%) were associated with pass rate observations, with an average of 9.5-9.7 observations per 15-year sampling window; similarly, 139-148 out of 300 ACBSP programs (46%-49%) were associated with pass rate observations, with an average of 6.2-6.3 observations per 15 year sampling window. In Table 1, the “Mean Relative Program Pass Rate” column presents the mean of the averages of the relative pass rate observations in each program’s 15-year sampling window. Depending on the section, the mean relative pass rate ranges from 0.818 to 0.869 for AACSB programs and from 0.742 to 0.797 for ACBSP programs. Thus, in our sample, both AACSB and ACBSP programs have average pass rates that are below the population mean rate for all programs. Similarly, the “Mean Program Candidates” column presents the mean of the averages of the number of first-time candidates sitting for the exam in each program’s 15-year sampling window. Depending on the section, the mean number of candidates ranges from 24.4 to 24.6 for AACSB programs and from 14.0 to 14.4 for ACBSP programs. Although our sample includes substantially more ACBSP than AACSB programs, AACSB programs generated more than twice as many pass rate observations and had almost twice as many first-time candidates sit for the exam.

Data Filter

As described above, according to the number of candidates in each program, we divided the data into mid-/large-size programs and small-size programs separately to reduce bias introduced by the small-cohort rate-cluster effect. Figure 2 suggests that this effect is moderated at a threshold of 15 to 20 candidates.

For mid-/large-size programs, when we filter the 504 program sample using a minimum threshold of 15 candidates in at least two years in each of the three subperiods, the number of programs surviving the filter is 72-74 depending on the exam section. When we increase the minimum threshold from 15 to 20 candidates, the number of programs surviving the filter falls from 72-74 to 41-42. The filter affects AACSB and ACBSP programs asymmetrically. Using the 15 candidate threshold, 14 of 300 ACBSP programs survive the filter compared to 58-60 of 204 AACSB programs that survive. Using the 20 candidate threshold, the number of programs surviving the filter falls to 5 ACBSP and 36-37 AACSB programs.

Figure 6 presents scatter plots showing the filter's effect on the AUD section. Panel A shows the relative pass rate distribution by number of program candidates for the full sample and Panel B presents this distribution for mid-/large-size programs that pass the 15 candidate filter. Scatter plots for other exam sections are similar. Relative to AACSB programs, very few ACBSP programs have more than 200 candidates sitting for the exam in a 5-year subperiod. Thus, programs with large candidate cohorts (more than 40 candidates per year on average) are concentrated in the AACSB group. Since the 20 candidate filter produced a sample with only 5 mid-/large-size programs in the ACBSP group, we elected to use the 15 candidate filter for the difference-in-differences analysis.

For small-size programs, we filter the sample using a maximum threshold of 15 candidates in at most one year in pre-accreditation subperiod. The number of programs surviving this filter is 91-95 depending on the exam section. AACSB and ACBSP programs are distributed symmetrically with 40-44 ACBSP programs and 51 AACSB programs.

Difference-In-Differences Findings

Table 2 and 3 present the results of the difference-in-differences analysis using the regression model in Eq. 15 and the 15 candidate filter for the mid-/large-size programs and small-size programs separately. These two tables present results for the composite pass rate and for each of the exam sections. To evaluate robustness, Table 2 and 3 also present the difference-in-differences analysis for programs with the 10 candidate filter. The 20 candidate filter, with only 5 surviving mid-/large- size programs in the ACBSP control group, resulted in a sample that was too small for statistical analysis.

In conventional null hypothesis significance testing, the p-value estimates the probability of observing a test statistic more extreme than the observed value given that the null hypothesis is true (Wasserstein & Lazar, 2016). Thus, the p-value can be regarded as a summary measure of the goodness-of-fit between the observations and the statistical assumptions that generated the null hypothesis. In a conventional statistical test, the observed p-value is compared to a specified significance level (typically 0.05 in social sciences) and the null rejected as unsupported by the data when the p-value is less than the specified significance level.

Over the last decade, this binary approach to significance testing has attracted substantial criticism (Greenland et al., 2016). For instance, the fact that the p-value is itself a random variable implies that two studies with 80% power, and observations selected at random from the same population, can arrive at opposite conclusions in as many as a third of the replications. From our perspective, the most important issue here is Gelman and Stern's (2006) related observation that the difference between two p-values that are close to 0.05, but lie on either side of 0.05, is often not a statistically significant difference even though the p-value < 0.05 result would be reported as significant and the p-value > 0.05 result as not significant. Thus, with this issue in mind, we have chosen to interpret the p-values from our regressions as a continuous measure of the evidence against the null and to avoid the binary decision framework (Dixon, 2003).

Mid-/Large-Size Programs

In Table 2, we obtain p-values that support the existence of a pre-treatment effect (β_s) but not a period effect (β_{t2}, β_{t3}) for the composite measure and most of four exam sections. Note that "pre-treatment effect" here references membership in the AACSB group and not the effect of AACSB accreditation relative to ACBSP accreditation. The p-values for the pre-treatment effects range between 0.027 and 0.129 for the 15 candidate model and between 0.010 and 0.062 for the 10 candidate model. The pre-treatment effect is strongest for the AUD and BEC exam sections and somewhat weaker for the FAR and REG sections. The magnitude of the effect is on the order of a 0.15 to 0.20 increase in the relative pass rate for programs in the AACSB group. On the other hand, the p-values for the phase-in period effects (β_{t2}) range between 0.862 and 0.980 for the 15 candidate model and between 0.799 and 0.983 for the 10 candidate model. The p-values for the post-accreditation period effects (β_{t3}) range between 0.670 and 0.996 for the 15 candidate model and between 0.304 and 0.962 for the 10 candidate model.

In the difference-in-differences analysis, the data do not support the existence of an AACSB accreditation effect relative to ACBSP accreditation. In this case, the p-values for the difference-in-differences coefficients (δ_2 and δ_3) range between 0.665 and 0.943 for the 15 candidate model and between 0.469 and 0.960 for the 10 candidate model.

Figure 7 shows the basis for this result. It presents a scatter plot of the AACSB and the ACBSP composite pass rates for the three accreditation subperiods. The AACSB and ACBSP subperiod pass rate means are also shown and are connected by dashed trend lines. The plots for the individual exam sections are similar to the composite pass rate plot in Figure 4 and, in all cases, we observe a similar pattern: (1) both the AACSB and ACBSP trend lines are essentially horizontal suggesting neither accreditation type has a detectable effect on the average pass rate; (2) the AACSB and ACBSP trend lines are parallel suggesting that the parallel trends assumption in the difference-in-differences analysis is satisfied; and, (3) the AACSB trend line lies above the ACBSP trend line indicating AACSB programs, on average, have higher pass rates than ACBSP programs.

In Table 2, we compare the 15 candidate difference-in-differences model with a 10 candidate model. In the latter, the number of programs that survive the filter increases to 78-80 AACSB and 17-18 ACBSP depending on the exam section. The coefficients and the coefficient p-values in the 10 candidate model are similar in magnitude to those in the 15 candidate model. The 10 candidate filter increases the sample size by about 30%. In this case, one would not expect significant changes in the magnitudes of the coefficients and p-values when the new observations are drawn from the same population which is the observed result. We interpret this as evidence for robustness (i.e., that the 15 candidate sample is representative of the underlying population).

Small-Size Programs

Interestingly, in Table 3, we obtain the p-values that support the existence of an AACSB accreditation effect relative to ACBSP accreditation (δ_2 and δ_3) for the composite measure and for all four exam sections among the small-size programs. In this case, the difference-in-differences estimate, δ_3, is constructed from the means of the pre- and post-accreditation periods, and has the p-values ranging between 0.000 and 0.001 for both 10 and 15 candidate models. It indicates that the influences of AACSB and ACBSP on the change in pass rate significantly differ over time between pre- and post-accreditation (i.e., 15 years) for all four exam sections. In addition, the difference-in-differences estimate, δ_2, is constructed from the means of the phase-in and post-accreditation periods. Its p-value ranges between 0.029 and 0.165 for the 15 candidate model (Composite: p = 0.029; AUD: p = 0.165; BEC: p = 0.048; FAR: p = 0.000; REG: p = 0.000) and between 0.007 and 0.071 for the 10 candidate model (Composite: p = 0.007; AUD: p = 0.026; BEC: p = 0.003; FAR: p = 0.021; REG: p = 0.071). It shows that, over time between phase-in and post-accreditation (i.e., 10 years), AACSB and ACBSP have significantly different influence on the change in pass rate in the BEC and FAR sections, however, these differences are somewhat weak for the AUD and REG exam sections.

Now it is important to exam which type of accreditation has more effect on the pass rate improvement as the small-size programs go through the process from pre-accreditation to phase-in stage to post-accreditation. According to Eq.15, at the phase-in period (subperiod 2), the pass rate in the AACSB group is $R_{st_2} = \beta_0 + \beta_s + \beta_{t_2} + \delta_{t_2}$, and the pass rate in the ACBSP group is $R_{st_2} = \beta_0 + \beta_{t_2}$. The difference in pass rate is $\Delta R_{st_2} = \beta_s + \delta_{t_2}$. As shown in Table 2, at the phase-in period, the AACSB group always has the higher pass rate than the ACBSP group for all four exam sections. Their difference ranges between 0.040 and 0.058 for the 15 candidate model (Composite: $\Delta R_{st_2} = \beta_s + \delta_{t_2} = 0.041$; AUD: $\Delta R_{st_2} = 0.043$; BEC: $\Delta R_{st_2} = 0.058$; FAR: $\Delta R_{st_2} = 0.040$; REG: $\Delta R_{st_2} = 0.045$) and between 0.095 and 0.159 for the 10 candidate model (Composite: $\Delta R_{st_2} = 0.118$; AUD: $\Delta R_{st_2} = 0.103$; BEC: $\Delta R_{st_2} = 0.159$; FAR: $\Delta R_{st_2} = 0.117$; REG: $\Delta R_{st_2} = 0.095$). Similarly, at the post-accreditation period (subperiod 3), the difference in pass rate between the AACSB and ACBSP groups is $\Delta R_{st_3} = \beta_s + \delta_{t_3}$. Results in Table 2 show that the AACSB group always has the higher pass rate than the ACBSP group for the 15 candidate model (Composite: $\Delta R_{st_3} = 0.160$; AUD: $\Delta R_{st_3} = 0.198$; BEC: $\Delta R_{st_3} = 0.165$; FAR: $\Delta R_{st_3} = 0.169$; REG: $\Delta R_{st_3} = 0.160$) and for the 10 candidate model (Composite: $\Delta R_{st_3} = 0.214$; AUD: $\Delta R_{st_3} = 0.254$; BEC: $\Delta R_{st_3} = 0.211$; FAR: $\Delta R_{st_3} = 0.246$; REG: $\Delta R_{st_3} = 0.231$). Clearly, the difference in pass rate between the ACBSP and ACBSP groups is increasing as the small programs go through their accreditation from the phase-in stage to the post-accreditation, that is, $\Delta R_{st_3} > \Delta R_{st_2}$, for both 10 and 15 candidate models in Table 2.

As an illustration of above results, the estimated composite pass rates using the regression model in Eq. 15 are presented in Figure 8, which is similar to the plots for the individual exam sections. In all cases, we observe a similar pattern: (1) both the AACSB and ACBSP trend lines are increasing, which suggests that both accreditation types have a positive effect on the pass rate; (2) at the phase-in and post-accreditation periods, the AACSB trend line lies above the ACBSP trend line indicating AACSB programs have higher pass rates than ACBSP programs; and, (3) the AACSB trend line increases more quickly than the ACBSP trend line, which indicates that AACSB accreditation has a greater impact on the pass rate of small-size programs compared to the ACBSP accreditation.

Conclusions

We examined the hypothesis that AACSB accreditation has a differential effect on the CPA exam pass rate relative to ACBSP accreditation using data for first-time candidates and a difference-in-differences design. This hypothesis was motivated by interviews we conducted with programs that had recently received AACSB accreditation. In the course of these interviews, we found that programs seeking AACSB accreditation frequently needed to add additional accounting faculty to meet AACSB accreditation standards and that AACSB criteria used to qualify faculty were generally more rigorous than that used by ACBSP.

Our analysis focused on programs that received initial accreditation from AACSB and ACBSP between 1990 and 2009. For each program, we collected fifteen years of pass rate observations divided into three subperiods with five years in each subperiod. The accreditation phase-in period consisted of the five years prior to and including the year the program received accreditation. The pre-accreditation period included the five years prior to the phase-in period and the post-accreditation period included the five years following the phase-in period. We used a difference-in-differences quasi-experimental design to compare the post-accreditation period with the pre-accreditation period.

This study contributes to the existing literature in several ways. First, although the current literature supports the correlation between AACSB accreditation and the CPA program pass rate, their causal inferences is still an open question. Our study lays some necessary experimental groundwork that advances the knowledge about the causal direction of this relationship. Second, AACSB and ACBSP memberships are two most popular options to consider when business programs seek for international accreditation. However, the literature pays little attention to the comparison between them, and most of existing research only investigate the program performance after accreditation has been granted. Considering the culture of continuous improvement required by both accreditations, it is necessary to examine the effects of AACSB and ACBSP during a period of time rather than at one point in time. To enrich this scant knowledge, this study tends to extend the current literature by examining the CPA exam pass rates before, during, and after accreditation to get a comprehensive understanding about the effect of AACSB and

ACBSP. In addition, as suggested by Jantzen and Pendleton (1994) that accreditation was related to the size of school, this study investigates the effects of accreditation on the CPA exam pass rate for the different size of programs, i.e., the number of candidates in each exam.

In our study of mid-/large-size programs receiving initial accreditation, we found membership in the AACSB group was associated with a significant pass rate effect. On average, membership in the AACSB group was associated with a 15% to 20% higher pass rate relative to the ACBSP group. This is consistent with the observations in the literature. In 2014, for example, the overall pass rate for first-time candidates from AACSB programs was 54.9%, compared to 42.7% for first-time candidates from ACBSP programs, and 42.8% for first-time candidates from IACBE programs (NASBA, 2014). However, we found no evidence that AACSB accreditation *per se* was causally related to changes in the exam pass rate. Thus, our results suggest that the observed consistent and stable difference in the pass rate between AACSB and ACBSP programs reflects systematic differences between programs that seek AACSB accreditation and programs that seek ACBSP accreditation. For example, in our sample, we observed that programs with large numbers of exam candidates (an average of 40 or more candidates per year) were almost always members of the AACSB group. This result is consistent with Boone et al. (2006) who found a weak relationship between AACSB accreditation and the CPA exam pass rate that could be explained by greater student selectivity.

For the small-size programs, we found that AACSB accreditation *per se* was causally related to changes in the exam pass rate relative to ACBSP accreditation. Both accreditations contribute to increased pass rate, however, AACSB membership had a greater increasing rate than ACBSP membership as the small-size programs went through the accreditation process from pre-accreditation to phase-in stage to post-accreditation. On average, membership in the AACSB group was associated with a 4% to 6% higher pass rate relative to the ACBSP group at the phase-in period, then became 16% to 20% higher at the post-accreditation period. Thus, our results suggest the increased difference in the pass rate between AACSB and ACBSP programs, which may reflect the differences between AACSB and ACBSP accreditation guidelines. For example, the AACSB puts greater emphasis on research and has high standards for the academically and professionally qualified faculty compared to the ACBSP (Brink & Smith, 2012). As a result, the AACSB programs are more likely to increase resources to faculty research and professional training, which brings more opportunities for their faculty to be aware of new advances in their discipline and discuss cutting-edge issues in classroom. Bline et al. (2016) observed that these practices improved teaching effectiveness and positively influenced candidate CPA exam performance. These effects may be stronger for the small-size programs which only had the limited resources before accreditation. As they seek AACSB accreditation which requires greater use of assessment data for planning (Hindi & Miller, 2000), small-size programs have more flexibility to change and grow faster by utilizing resources offered, such as adjusting content delivery and offering a variety of delivery formats. It is consistent with Hindi and Miller (2000) who found accounting departments with AACSB accreditation identified life-long learning as a skill to assess to a significantly greater extent than did those departments with other accreditations.

The results of our study should be interpreted in light of the following limitations. Since the exam's composition, administration, and content change periodically over the study period, and programs respond to these changes with changes in program emphasis and curriculum, program pass rates should be interpreted as outcomes of a dynamic adaptive process in which programs evolve in response to an evolving exam regime. Similarly, the standards used for accreditation evolve in a similar way and should be considered part of this process. We adjust for this dynamic process in two respects. First, we rescale the pass rate relative to the mean population pass rate which adjusts for systematic variation in the exam regime. Second, since pass rate observations are assigned to the time index relative to accreditation year rather than to the calendar year in which the observation occurred, the exam regime can be regarded as randomly assigned which implies that pass rate variation induced by the exam regime averages out in our study design. However, it should be noted that these adjustments, although helpful, are imperfect and their precise effect on the measurement issue unknown.

Second, Ryan et al. (2014) used Monte Carlo methods to study the effect of specification choice on the accuracy of difference-in-differences estimates. They found that parameter estimates were sensitive to model specification when the probability of treatment was correlated with pre-intervention levels and trends. Since AACSB accreditation choice appears to be related to pre-accreditation CPA exam pass rates, we suggest that future research consider the use of techniques like propensity score matching (Shipman et al., 2017), which uses observed characteristics to predict the probability of group membership, to generate the counterfactual treatment group.

Third, we chose the ACBSP programs as the control group in the difference-in-differences analysis. In order to better understand the influence of AACSB accreditation on the CPA exam pass rate, additional research, involving non-accredited programs, will be necessary. Another limitation involves the small number of observations for the ACBSP group, especially, the sample size of ACBSP programs with an intermediate to a large number of candidates in this study. The ACBSP was founded in 1988, and our data sample consists of programs that received initial accreditation between 1990 and 2009. So, the ACBSP was relatively young without sufficient accredited programs during the period of interest. In order to test the generalizability of results, an additional analysis should be conducted by including the more recent accredited programs from AACSB or ACBSP. Finally, the current research compares the CPA exam pass rate among various stages of the accreditation process (i.e., pre-accreditation, phase-in period, and post-accreditation) without controlling for the influence of program characteristics, such as the enrollment of students or the quality of students. It would be interesting to investigate how students differ between AACSB and ACBSP accreditations and exam their influence on the CPA exam pass rate.[1]

[1] **Acknowledge:** We would like to thank the National Association of State Boards of Accountancy for the support received, as the recipient of a 2015 Accounting Education Research Grant, in the conduct and preparation of this research. We also would like to thank the Editor, Associate Editor, and two anonymous referees for their many useful comments.
Disclosure statement: The authors report there are no competing interests to declare.

References

Albring, S.M., & Elder, R.J. (2020). Research initiatives in accounting education: Managing academic programs. *Issues in Accounting Education*, *35*(4), 61-74.

Angrist, J. D., & Pischke, J-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton, N. J.

Angrist, J. D., & Pischke, J-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, *24*(2), 3-30.

Barilla, A., Jackson, R., & Mooney J. (2008). The CPA exam as a postcurriculum accreditation assessment. *Journal of Education for Business, 83*(5), 270-274.

Bergin, J., Morgan, J., & Sallee, L. (2011). AACSB accreditation, ACBSP accreditation and CPA exam success rates. *Midwestern Business and Economic Review, 47*, 7-15.

Bline, D. M., Perreault, S., & Zheng, X. (2016). Do accounting faculty characteristics impact CPA exam performance? An investigation of nearly 700,000 examinations. *Issues In Accounting Education, 31*(3), 291-300.

Boleslavsky, R., & Cotton, C. (2015). Grading standards and education quality. American Economic Journal: Microeconomics, *7*(2), 248–279.

Boone, J., Legoria, J., Siefert, D., & Stammerjohan, W. (2006). The associations between accounting program attributes, 150-hour status and CPA exam pass rates. *Journal of Accounting Education*, *24*(4), 202-215.

Brink, K. & Smith, C. (2012). A comparison of AACSB, ACBSP, and IACBE accredited U.S. business programs: An institutional resource perspective. *Business Education and Accreditation*, *4*(2), 1-16.

Bunker, R., Cagle, C., & Harris, D. (2014). Comparison of AACSB accredited and ACBSP accredited institutions using the CPA exam as a post-curriculum assessment. *Journal of Accounting & Finance*, *14*(6), 127-132.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39–67.

Dixon, P. (2003). The p-value fallacy and how to avoid it. *Canadian Journal of Experimental Psychology*, *57*(3), 189-202.

Franklin, M. A., & Myers, J. K. (2016). Enhancing CPA exam pass rates for second career students. *Journal of Business and Educational Leadership, 6*(1), 27-36.

Fuller, L.R., & Hargadon, J.M. (2008). MBA, CMA, and CPA: Natural partners in the 150-hour requirement. *Management Accounting Quarterly*, *9*(2), 39-49.

Gaynor, G., Korb, P., Gerlowski, D., & Zhang, T. (2019). An alternate state in mind: The effect of CPA exam credit-hour requirements and economic competitiveness on state-level exam candidate pools and pass rates. *Accounting Education*, *28*(6), 621-641.

Gelman, A., & Stern, H. (2006). The difference between “significant”' and “not significant”' is not itself statistically significant. *American Statistician*, *60*(4), 328-331.

Grant, C., Ciccotello, C., & Dickie, M. (2002). Barriers to professional entry: How effective is the 150-hour rule? *Journal of Accounting and Public Policy*, *21*(1), 71-93.

Greenland, S., Senn S.J., Rothman K.J., Carlin, J.B., Poole, C., Goodman, S.N., & Altman, D.G. (2016). Statistical tests, p values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, *31*(4), 337-350.

Hindi, N., & Miller, D. (2000). A survey of assessment practices in accounting departments of colleges and universities. *Journal of Education for Business*, *75*(5), 286-290.

Jantzen, R. H., & Pendleton, T. A. (1994). Preferences of the American assembly of collegiate schools of business. *Journal of Education for Business*, *70*(1), 1-10.

Kundu, G.K. (2020). Modeling the critical barriers to implementation of AACSB accreditation in business schools. *Benchmarking: An International Journal*, *27*(2), 550-570.

Nagle, B.M., Menk, K.B., & Rau, S.E. (2018). Which accounting program characteristics contribute to CPA exam success? A study of institutional factors and graduate education. *Journal of Accounting Education, 45*, 20-31.

NASBA (National Association of State Boards of Accountancy). (1995). *Candidate Performance on the Uniform CPA Exam.* 1995 Edition, Nashville, TN.

NASBA (National Association of State Boards of Accountancy). (2005). *Candidate Performance on the Uniform CPA Exam.* 2005 Edition Revised, Nashville, TN.

NASBA (National Association of State Boards of Accountancy). (1985-2004). *Candidate Performance on the Uniform CPA Exam.* Nashville, TN.

Marts, J. A., Baker, J. D., & Garris, J. M. (1988). Success on the CPA examination in AACSB accredited and non-accredited schools. *Accounting Educators' Journal*, *1*, 97–105.

Morgan, J., Bergin, J., & Sallee, L. (2008). An investigation of the relationship between AACSB business school accreditation and CPA exam success rates. *Journal of Business and Leadership*, *4*(2), 20-32.

Morgan, J., Bergin, L. J., & Sallee, L. (2012). Three types of business school accreditation and their relationships to CPA exam scores of graduates. *Advances in Business Research*, *3*(1), 25-35.

Panhans, M.T., & Singleton, J.D. (2017). The empirical economist's toolkit: From models to methods. *History of Political Economy*, *49*(Suppl.), 127-157.

Roller, R., Andrews, B., & Bovee, S. (2003). Specialized accreditation of business schools: A comparison of alternative costs, benefits and motivations. *Journal of Education for Business*, *78*(4), 197-204.

Ryan, A.M., Burgess, Jr., J.F., & Dimick, J.B. (2015). Why we should not be indifferent to specification choices for difference-in-differences. *Health Services Research*, *50*(4), 1211-1235.

Shipman, J.E., Swanquist, Q.T., & Whited, R.L. (2017). Propensity score matching in accounting research. *The Accounting Review*, *97*(1), 213-244.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: Context, process, and purpose. *The American Statistician*, *70*(2), 129-133.

Table 1: Data Sample Characteristics

Accreditation Type	Programs in Sample	Section	Programs With Pass Rate Observations	Number of Observations	Mean Observations Per Program	Mean Relative Program Pass Rate*	Mean Program Candidates*
AACSB	205	AUD	193	1861	9.6	0.869 (0.349)	24.4 (22.7)
		BEC	192	1871	9.7	0.835 (0.339)	24.6 (22.6)
		FAR	191	1830	9.6	0.818 (0.371)	24.6 (22.7)
		REG	193	1842	9.5	0.848 (0.359)	24.4 (22.6)
ACBSP	301	AUD	145	896	6.2	0.779 (0.421)	14.1 (22.0)
		BEC	148	924	6.2	0.742 (0.413)	14.0 (21.8)
		FAR	139	872	6.3	0.743 (0.416)	14.4 (22.5)
		REG	142	882	6.2	0.797 (0.429)	14.3 (22.2)

* Standard deviations shown in parenthesis.

Table 1: Presents the overall characteristics of the data sample. The sample consisted of pass rate and first-time candidate count observations from the 1985-2014 period for 504 programs that received business school accreditation between 1990 and 2009. The *Mean Relative Program Pass Rate* column presents the mean of the averages of the relative pass rate observations in each program’s 15-year sampling window. The *Mean Program Candidates* column presents the mean of the averages of the number of first-time candidates sitting for the exam in each program’s 15-year sampling window. Although the sample included substantially more ACBSP programs, AACSB programs generated twice as many observations as ACBSP programs and, on average, AACSB programs had 10% higher relative pass rates and 73% more candidates sitting for the exam.

Table 2: Difference-In-Differences Analysis (Mid-/Large-Size Program)

Exam Section	Program Filter: 15 Candidates Per Year in at least 2 Years Per 5-Year Subperiod			Program Filter: 10 Candidates Per Year in at least 2 Years Per 5-Year Subperiod		
	Surviving Programs	**Parameter Estimate**	**P-Value**	**Surviving Programs**	**Parameter Estimate**	**P-Value**
Composite						
Constant (β_0)	57 AACSB	0.671**	0.000	75 AACSB	0.649**	0.000
Treatment (β_s)	14 ACBSP	0.180*	0.046	17 ACBSP	0.187*	0.020
Period 2 (β_{t2})		0.003	0.980		0.026	0.799
Period 3 (β_{t3})		0.001	0.996		0.048	0.636
DID Period 2 (δ_2)		-0.009	0.943		-0.039	0.731
DID Period 3 (δ_3)		-0.039	0.760		-0.051	0.652
AUD						
Constant (β_0)	61 AACSB	0.644**	0.000	80 AACSB	0.612**	0.000
Treatment (β_s)	14 ACBSP	0.200*	0.034	17 ACBSP	0.215**	0.010
Period 2 (β_{t2})		-0.016	0.894		0.016	0.882
Period 3 (β_{t3})		0.026	0.827		0.109	0.304
DID Period 2 (δ_2)		0.024	0.855		-0.018	0.875
DID Period 3 (δ_3)		-0.047	0.721		-0.084	0.469
BEC						
Constant (β_0)	60 AACSB	0.664**	0.000	80 AACSB	0.680**	0.000
Treatment (β_s)	14 ACBSP	0.201*	0.027	18 ACBSP	0.167*	0.037
Period 2 (β_{t2})		0.020	0.862		-0.002	0.983
Period 3 (β_{t3})		-0.015	0.895		0.005	0.962
DID Period 2 (δ_2)		-0.041	0.750		-0.012	0.919
DID Period 3 (δ_3)		-0.044	0.730		-0.028	0.807
FAR						
Constant (β_0)	58 AACSB	0.666**	0.000	79 AACSB	0.647**	0.000
Treatment (β_s)	14 ACBSP	0.164	0.080	18 ACBSP	0.154	0.062
Period 2 (β_{t2})		-0.013	0.910		0.018	0.864
Period 3 (β_{t3})		-0.051	0.670		-0.028	0.790
DID Period 2 (δ_2)		-0.044	0.741		-0.057	0.626
DID Period 3 (δ_3)		-0.019	0.887		0.006	0.960
REG						
Constant (β_0)	60 AACSB	0.679**	0.000	78 AACSB	0.648**	0.000
Treatment (β_s)	14 ACBSP	0.151	0.129	17 ACBSP	0.169	0.056
Period 2 (β_{t2})		0.004	0.973		0.027	0.813
Period 3 (β_{t3})		0.034	0.789		0.073	0.520
DID Period 2 (δ_2)		0.015	0.913		-0.018	0.883
DID Period 3 (δ_3)		-0.061	0.665		-0.045	0.720

Table 2: Presents difference-in-differences regression results for the composite pass rate and by individual section for programs that pass the 15 candidate and 10 candidate thresholds in at least two years in each of the three subperiods. ‘*’ and ‘**’ indicate significance levels of 0.05 and 0.01 respectively.

Table 3: Difference-In-Differences Analysis (Small-Size Program)

Exam Section	Program Filter: 15 Candidates Per Year in at most 1 Year of Pre-Accreditation Subperiod			Program Filter: 10 Candidates Per Year in at most 1 Year of Pre-Accreditation Subperiod		
	Surviving Programs	Parameter Estimate	P-Value	Surviving Programs	Parameter Estimate	P-Value
Composite						
Constant (β_0)	51 AACSB	0.405**	0.000	33 AACSB	0.345**	0.000
Treatment (β_s)	40 ACBSP	-0.113*	0.023	28 ACBSP	-0.119	0.054
Period 2 (β_{t2})		0.081	0.122		0.118	0.066
Period 3 (β_{t3})		0.161**	0.002		0.214**	0.001
DID Period 2 (δ_2)		0.154*	0.029		0.237**	0.007
DID Period 3 (δ_3)		0.273**	0.000		0.333**	0.000
AUD						
Constant (β_0)	51 AACSB	0.356**	0.000	33 AACSB	0.326**	0.000
Treatment (β_s)	44 ACBSP	-0.063	0.242	31 ACBSP	-0.108	0.108
Period 2 (β_{t2})		0.112*	0.045		0.134*	0.048
Period 3 (β_{t3})		0.154**	0.006		0.187**	0.006
DID Period 2 (δ_2)		0.106	0.165		0.211*	0.026
DID Period 3 (δ_3)		0.261**	0.001		0.362**	0.000
BEC						
Constant (β_0)	51 AACSB	0.349**	0.000	33 AACSB	0.292**	0.000
Treatment (β_s)	43 ACBSP	-0.083	0.100	31 ACBSP	-0.093	0.123
Period 2 (β_{t2})		0.076	0.146		0.087	0.157
Period 3 (β_{t3})		0.171**	0.001		0.238**	0.000
DID Period 2 (δ_2)		0.141*	0.048		0.252**	0.003
DID Period 3 (δ_3)		0.248**	0.001		0.304**	0.000
FAR						
Constant (β_0)	51 AACSB	0.358**	0.000	33 AACSB	0.300**	0.000
Treatment (β_s)	42 ACBSP	-0.110*	0.033	29 ACBSP	-0.093	0.148
Period 2 (β_{t2})		0.048	0.378		0.082	0.219
Period 3 (β_{t3})		0.112*	0.039		0.154*	0.021
DID Period 2 (δ_2)		0.150*	0.040		0.210*	0.021
DID Period 3 (δ_3)		0.279**	0.000		0.339**	0.000
REG						
Constant (β_0)	51 AACSB	0.375**	0.000	33 AACSB	0.293**	0.000
Treatment (β_s)	43 ACBSP	-0.108*	0.043	30 ACBSP	-0.072	0.274
Period 2 (β_{t2})		0.080	0.150		0.166*	0.013
Period 3 (β_{t3})		0.137*	0.013		0.203**	0.003
DID Period 2 (δ_2)		0.153*	0.042		0.167	0.071
DID Period 3 (δ_3)		0.268**	0.000		0.303**	0.001

Table 3: Presents difference-in-differences regression results for the composite pass rate and by individual section for programs that pass the 15 candidate and 10 candidate thresholds in at most one year in the pre-accreditation subperiod. '*' and '**' indicate significance levels of 0.05 and 0.01 respectively.

Figure 1: Average CPA Exam Section Pass Rates 1985-2015 For First-Time Candidates

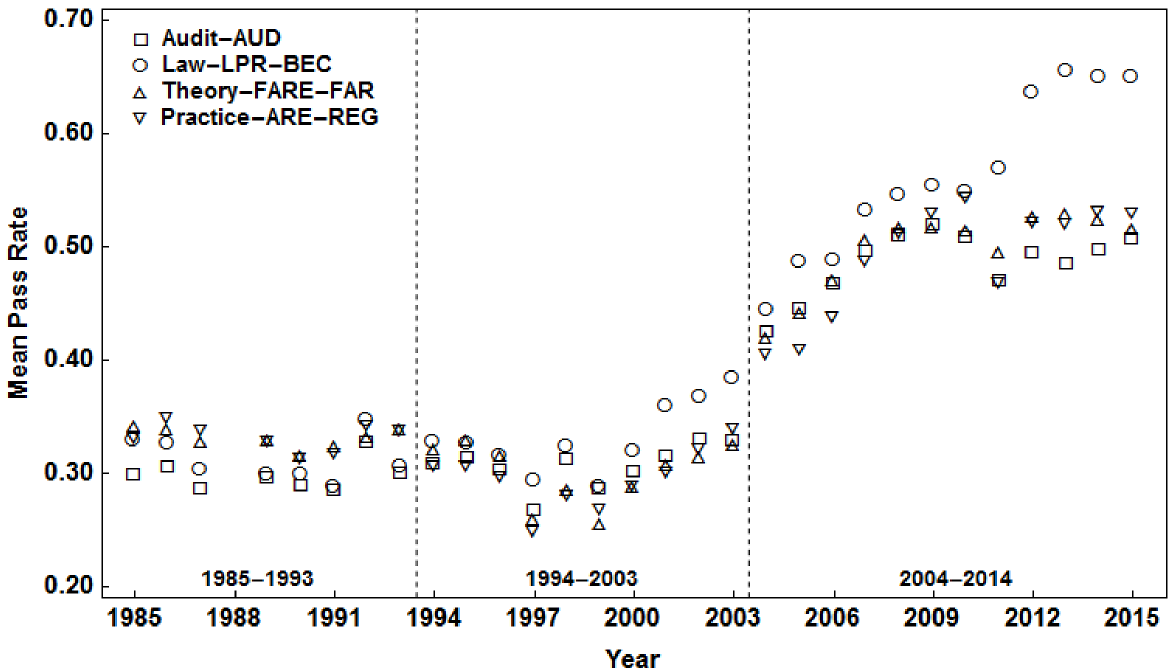
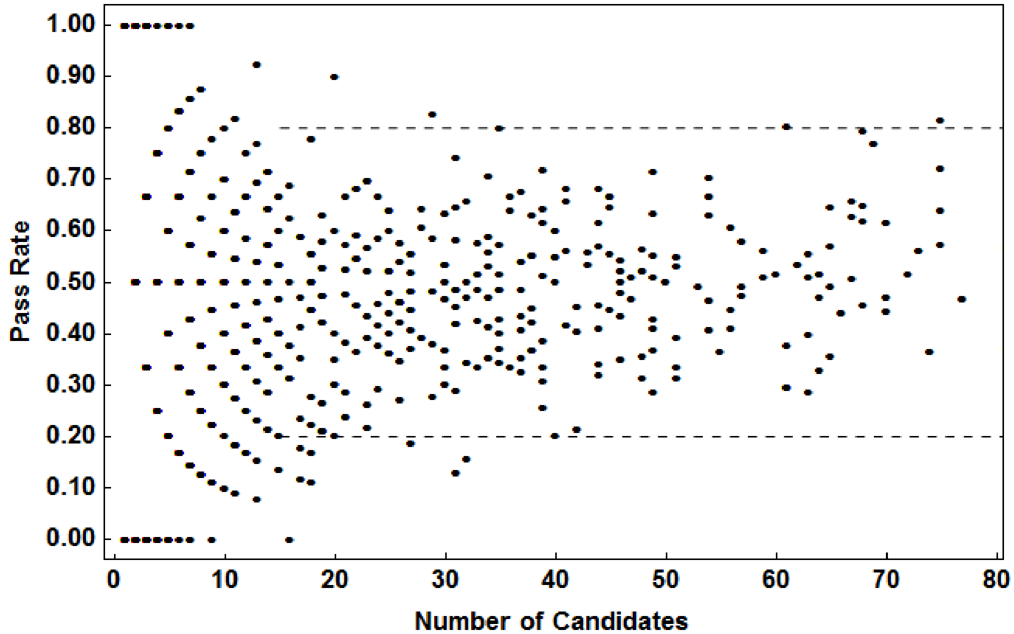


Figure 1: Presents average annual CPA exam pass rates by section for first-time candidates during the 1985-2015 period. No data were available for the 1988 exam administrations. Between 1985 and 2003, pass rates were reported separately for the May and November administrations. The average annual pass rates for this period are weighted averages of the May and November pass rates using the number of candidates as weights. The vertical dashed lines mark discontinuities due to major revisions to the exam in 1994 and 2004.

Figure 2: Distribution of CPA Exam Program Pass Rates (2014 AUD Section)
Panel A: By Number of Candidates



Panel B: By Average Candidate Score

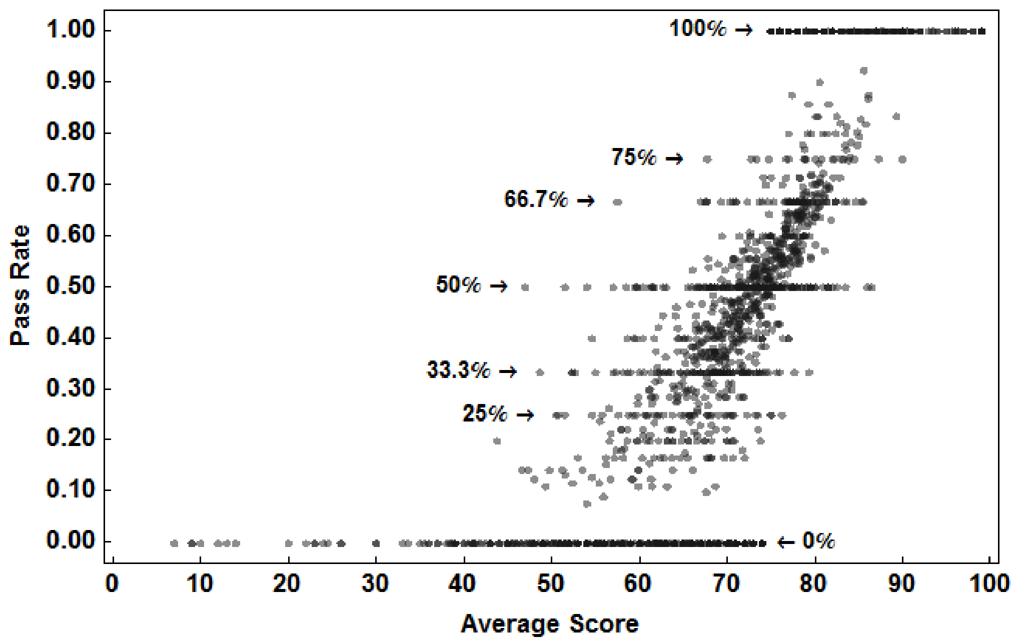


Figure 2: Presents the distribution of program pass rates for first-time candidates sitting for the AUD section of the exam in 2014. Panel A presents the distribution of pass rates by the number of program candidates and shows that programs with small numbers of candidates introduce variation bias by inflating the interval between pass rate observations. For example, when only a single candidate sits for the exam in a given year, the only possible pass rates are 0% and 100%. When two candidates take the exam, the possible pass rates are 0%, 50%, and 100%, and so on. The horizontal dashed lines indicate that this bias is reduced to background levels at about 15 candidates. Panel B presents the distribution of program pass rates by average score on the exam and shows that the pass rate "attractors" for programs with small numbers of candidates in Panel A are associated with an abnormally wide range of average scores.

Figure 3: Composite Pass Rate Construction

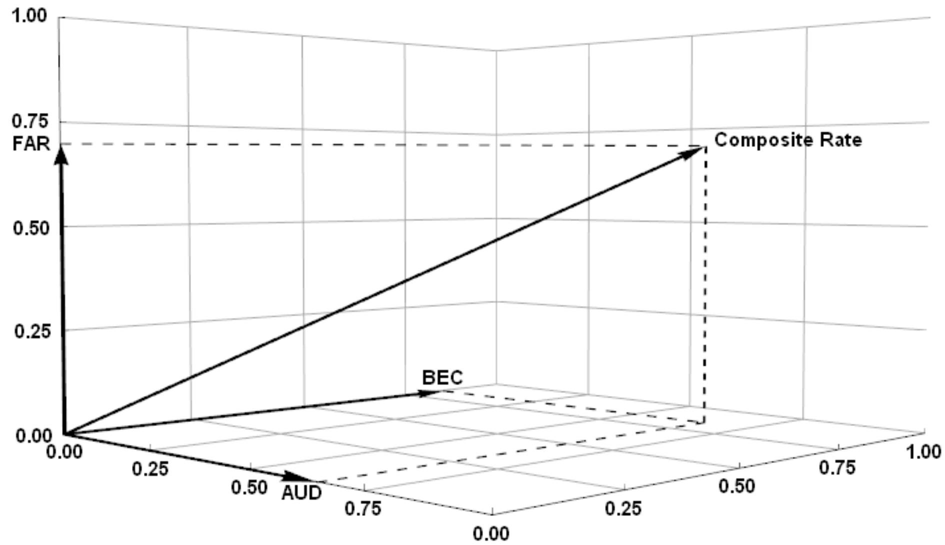
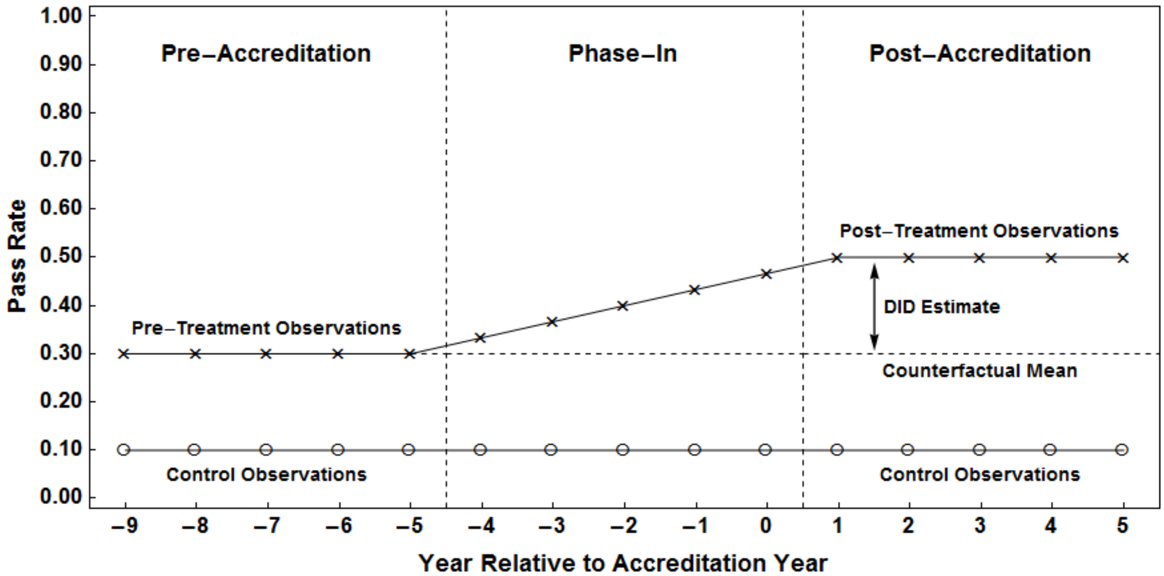


Figure 3: Illustrates the distance metric used to construct a composite pass rate. The composite program pass rate R_i^* is the equally-weighted 2nd-order mean of the section pass rates which, in geometric terms, is the Euclidean length of the vector constructed from these rates. For example, assume the exam consists of three sections (since graphical illustrations are limited to at most three spatial dimensions) and program i's observed pass rates are: $R_{i\text{AUD}}$= 0.65, $R_{i\text{BEC}}$= 0.85, and $R_{i\text{FAR}}$= 0.70. If these pass rates measure performance along orthogonal dimensions in the outcome space, then the Euclidean length of the vector constructed from the section pass rates is $\sqrt{R_{iAUD}^2 + R_{iBEC}^2 + R_{iFAR}^2} = 1.279$ which is the distance between the origin at [0, 0, 0] and the composite pass rate at [0.65, 0.85, 0.70]. To facilitate comparison with the section pass rates, which lie in the zero to one interval, normalize the Euclidean distance to unit length to obtain $R_i^* = 1.279/\sqrt{3} = 0.738$; thus, program i has a composite pass rate that is 73.8% of the rate for a program with 100% pass rates on all sections. Since our study uses rescaled pass rates, rather than absolute rates, the interpretation of the composite pass rate would be made relative to the population mean pass rate.

Figure 4: The Difference-In-Differences Design

Panel A: Study Design



Panel B: Parallel Trends Assumption

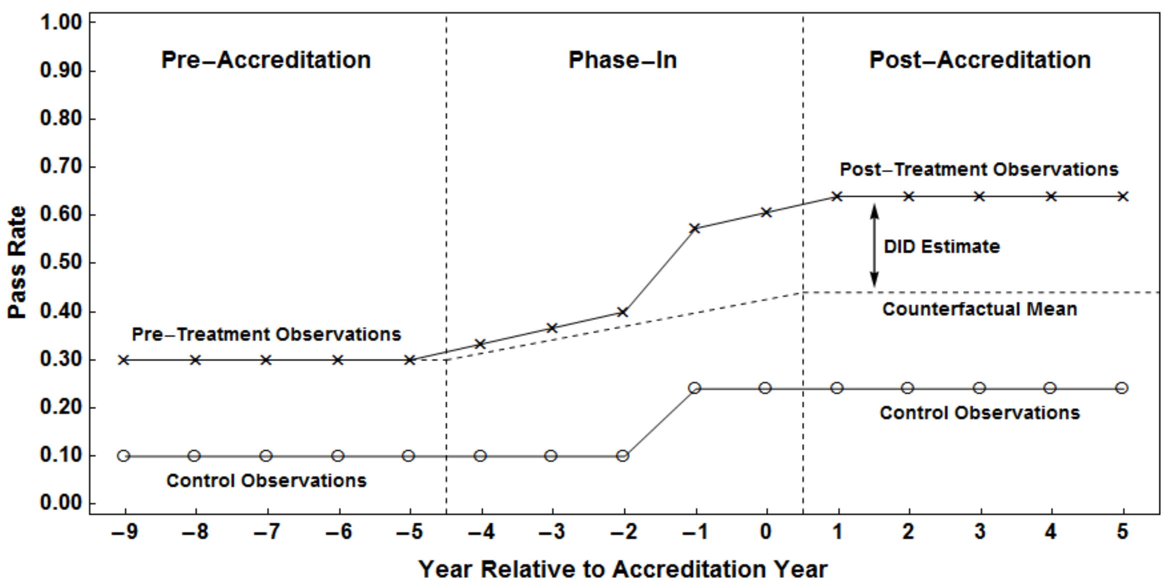


Figure 4: Illustrates the difference-in-differences design used in the study. The analysis is structured as an event study in which the event (accreditation) occurs in Year 0. Annual pass rate observations from the program group experiencing the treatment (AACSB accreditation) are designated using x's and observations from the control group (ACBSP accreditation) are designated using circles. The observations cover a 15-year period divided into three 5-year subperiods and, for simplicity, the figure does not show random variation in the observations. The 10 years of pass rate observations prior to, and including, the accreditation year are divided into a 5-year pre-accreditation period and a 5-year phase-in period to account for changes made to prepare for accreditation. The 5 years of pass rate observations following the accreditation year are the post-accreditation observations. In Panel A, the difference-in-differences (DID) estimator compares the change in the average subperiod pass rate in the treatment group to the change in the control group. In terms of levels, this is equivalent to comparing the post-accreditation pass rate in the

treatment group to its estimated counterfactual value given that the treatment group did not receive the treatment. The estimated counterfactual pass rate is constructed on the parallel trends assumption that, in the absence of the treatment, the treatment group would experience the same pass rate trend as the control group. Panel B illustrates the role of the parallel trends assumption. In Panel B, the programs experience a systematic shock to the pass rate in the year prior to accreditation. Since the shock affects both groups similarly, it affects the level of the pass rates but not the relative change in the pass rates which remains the same in both panels.

Figure 5: Distribution of Program Accreditations by Year and Type

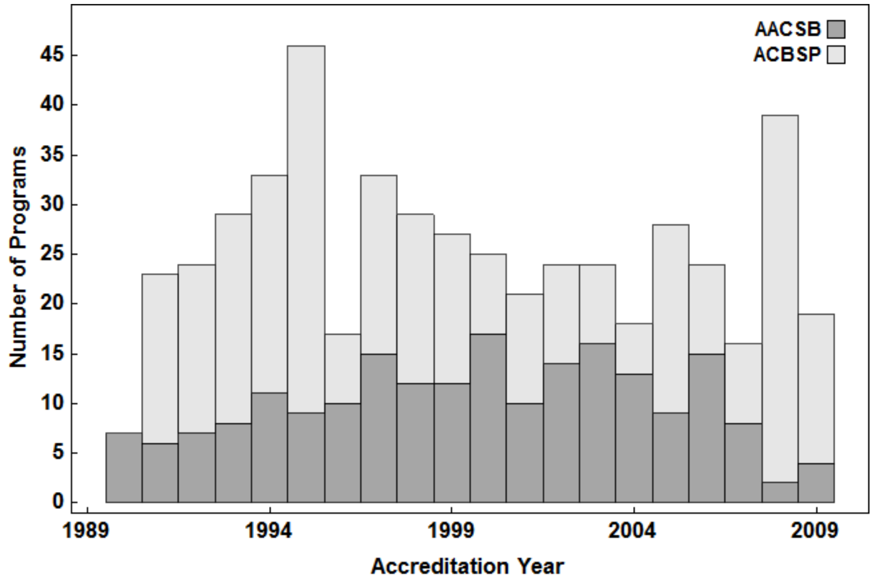
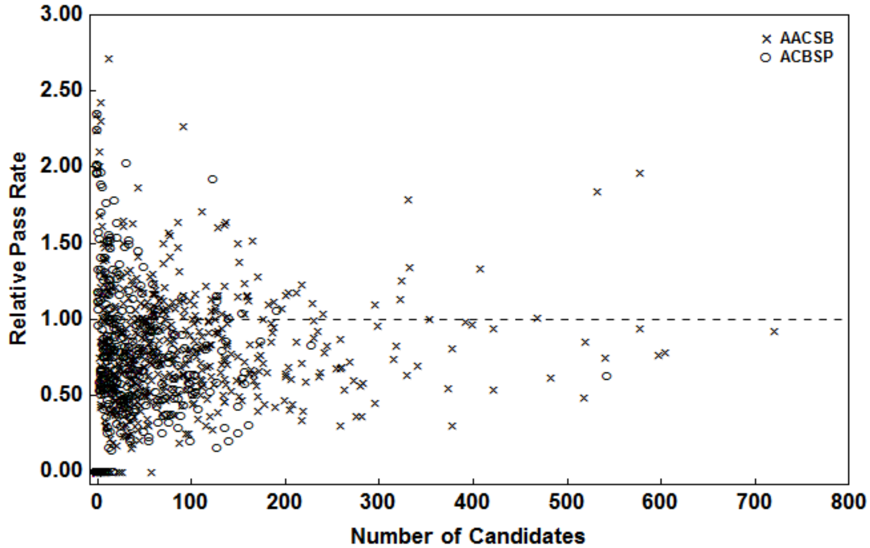


Figure 5: Presents the distribution of AACSB and ACBSP accreditations by year for the sample of 504 (204 AACSB and 300 ACBSP) programs.

Figure 6: Distribution of Relative Pass Rates by Number of Candidates (AUD Section)

Panel A: Full Sample



Panel B: Programs Passing a 15 Candidate Filter

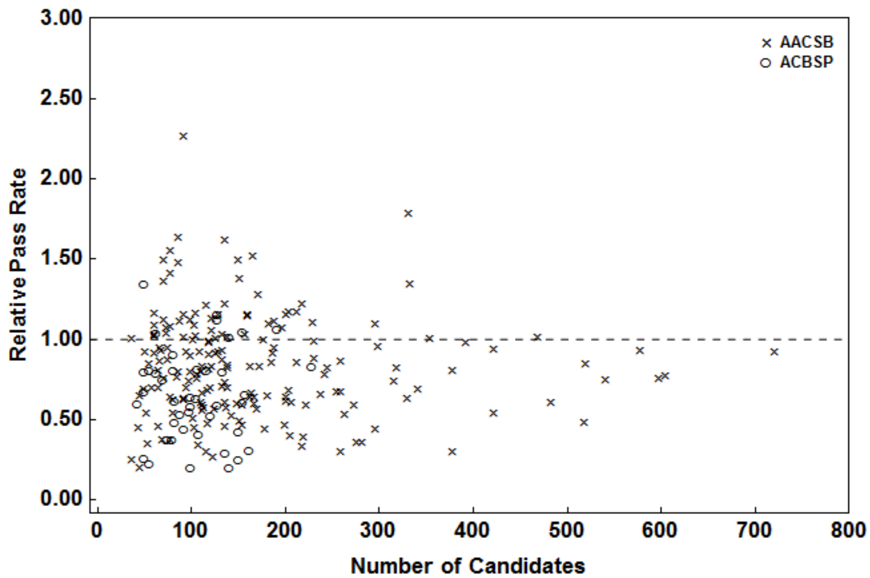


Figure 6: Presents the distribution of pass rates by number of candidates for the sample of 504 programs. The figure presents the results for the AUD section of the exam but scatter plots for the other sections are similar. Each marker identifies a program's relative pass rate and number of exam candidates for one of the three 5-year subperiods. Panel A shows the distribution for the full sample of 504 programs (204 AACSB and 300 ACBSP). Programs that produce large numbers of candidates (200 candidates in a 5-year subperiod) are overrepresented in the AACSB group. Panel B shows the distribution for programs that pass the 15 candidate threshold in at least two years in each of the three subperiods (61 AASCB and 14 ACBSP).

Figure 7: Distribution of Relative Pass Rates by Subperiod and Accreditation Type for Mid-/Large-Size Programs Surviving the 15 Candidate Filter

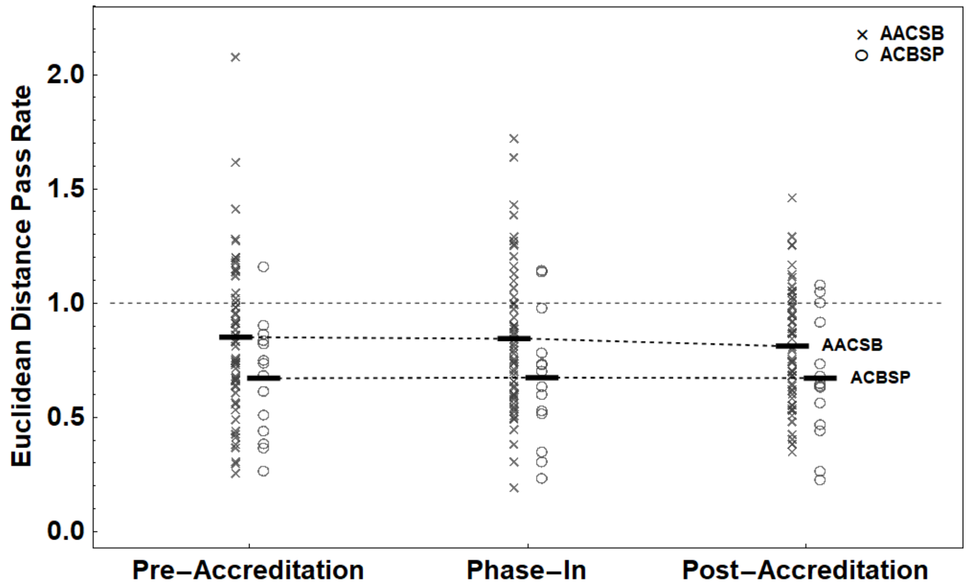


Figure 7: Presents the distribution of relative pass rates by subperiod for mid-/large-size programs that pass the 15 candidate threshold in at least two years in each of the three subperiods. The figure presents the results for the Euclidean distance composite pass rate but plots for the individual exam sections are similar. Dashed trend lines connect the subperiod means for the AACSB group and the ACBSP group. The dashed line at a relative pass rate of 1.0 is the population mean pass rate.

Figure 8: Relative Pass Rate Trends by Using Difference-In-Differences Analysis for Small-Size Programs Following the 15 Candidate Filter

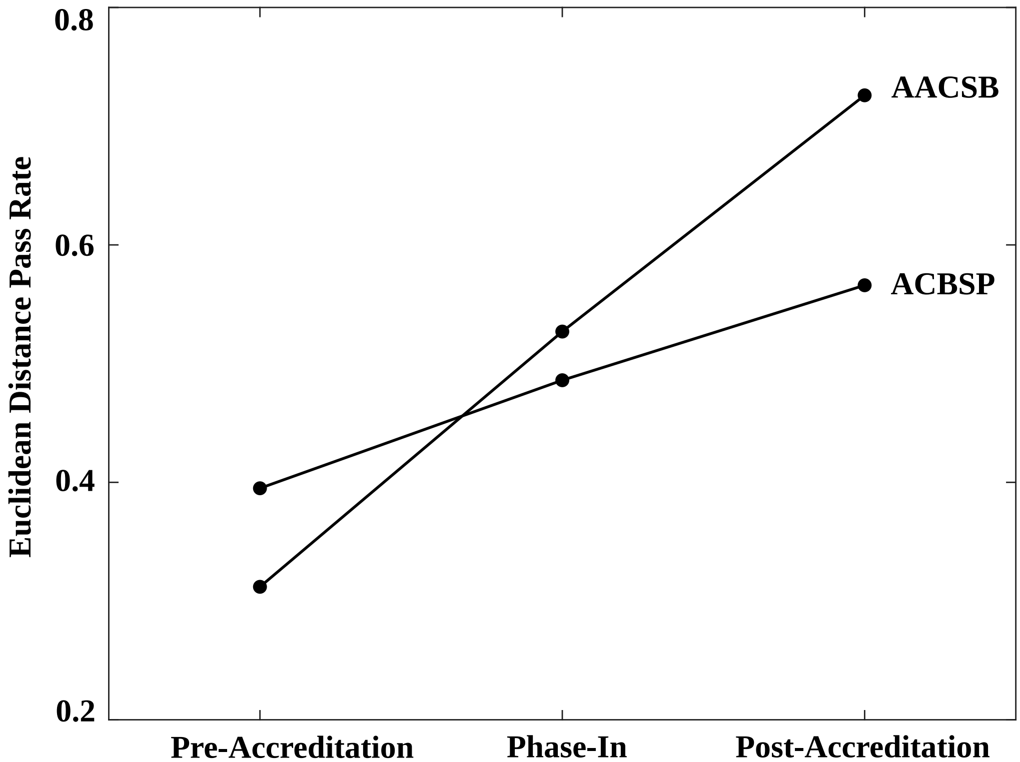


Figure 8: Presents the relative pass rate trends by using difference-in-differences analysis for small-size programs that satisfy the 15 candidate threshold in at most one year of pre-accreditation subperiod. The figure presents the results for the Euclidean distance composite pass rate but plots for the individual exam sections are similar. Solid lines connect the estimated pass rate at each subperiod by using the Eq. 15 for the AACSB group and the ACBSP group.