Opportunity, ethnicity, gender, and CPA exam performance

Arya Espahbodi
Faculty in Mathematics
Bard High School Early College Baltimore

Linda Espahbodi
President and CEO
Inspiring Accounting Talent for a Sustainable Society

Reza Espahbodi*
Dibble Professor of Accounting, School of Business
Washburn University

Rosemary Walker
Professor of Economics, School of Business
Washburn University

G. Thomas White
Professor Emeritus
College of William and Mary

* Corresponding author

Please do not quote or reference without authors' consent.

Acknowledgments: We appreciate the suggestions and guidance from Jeanette Franzel and Frank Wood. Also appreciated are the helpful comments from the participants at the 2018 Kaw Valley Bank Research Seminar Series at Washburn University. We thank Maureen Dombrow for data cleaning assistance on this project. The remaining errors are ours.

Reza Espahbodi and G. Thomas White acknowledge financial support from National Association of State Boards of Accountancy (NASBA); and Reza Espahbodi and Rosemary Walker acknowledge financial support from Washburn University.

Opportunity, ethnicity, gender, and CPA exam performance

ABSTRACT

Given the preeminence of the CPA certification as a measure of professional achievement and a critical element to advancement in the profession, as well as the concerns over lack of diversity in the accounting profession (AICPA 2017), a key policy question is how to improve candidates' performance on the CPA exam. In this paper, we examine the role of educational and environmental (socioeconomic and segregation) factors representing opportunity, as well as gender and ethnicity (as defined by the National Association of State Boards of Accountancy), on the CPA exam performance. To accomplish this, we first document CPA exam performance across various demographic, educational, and environmental factors. We then develop several multivariate models to understand the influence of various educational and environmental factors representing opportunity on the CPA exam performance of these groups. Finally, we springboard from our findings to offer suggestions to educators, professional firms, and CPA societies, to implement new, or modify current, programs to meet the profession's need for more qualified CPAs and its diversity/inclusion goals.

Keywords: CPA exam, socioeconomic status, community segregation, opportunity

Opportunity, ethnicity, gender, and CPA exam performance

Accounting firms must keep up with rapid changes in technology and business environment to competitively serve their clients, and are eager to hire more qualified and diverse college graduates (AICPA 2017). Because the CPA certification is a measure of knowledge and professional achievement and a critical element to advancement in the profession, accounting firms are especially interested in candidates who plan on getting to and passing the CPA exam, with an eye on achieving their diversity/inclusion goals. In fact, some firms even offer a generous bonus to new hires who pass the exam within the first two years of employment. On the supply side, accounting firms are excellent places for college graduates to launch their business careers. This type of environment allows for individuals to partake in life-long learning and thus, reap the reward of multiple career opportunities during their lifetime. The key that opens the door to these opportunities is successful completion of the CPA exam, leading to licensure as a CPA.

Given the demand for, and private returns to, CPA attainment, a key policy question is how to improve candidates' performance on the exam. CPA exam pass rate is lower than that in other professional exams, e.g., medical, and lowest for African-Americans and Hispanics and to a lesser extent for females. Only 44 percent of CPA candidates overall, and 20, 32, 41 percent of African-American, Hispanic, and female candidates passed the CPA exam over 2005-2016. These rates concern the profession.

While prior research (e.g., Bline, Perreault, and Zheng 2016a; Booker 2005; Morgan 2015; Trinkle, Scheiner, Baldwin, and Krull 2016) has investigated the association between ethnicity, gender, and select university and candidate characteristics, none has investigated the influence of socioeconomic condition, community segregation, and other opportunity factors (collectively referred to as opportunity) on the CPA exam performance. Evidence-informed policy recommendations/decisions should consider all significant factors potentially affecting performance on the CPA exam. This study, therefore, examines the relation between opportunity, ethnicity, and gender and performance on the CPA exam. Opportunity in this context encompasses variables such as: (a) availability of AP courses at high school, (b) access to counselors at high-school, (c) quality of accounting related coursework and relevance to the CPA exam, (d) exposure to mentors among accounting faculty or those already employed in the accounting profession, (e) relevant internship or some form of work experience with an accounting firm, and (f) free access to CPA exam review courses through future employer.[1]

Our study is inspired by the work of Sean Reardon and his colleagues at Stanford and Harvard (e.g., Fahle and Reardon *forthcoming*; Reardon 2016; Reardon, Kalogrides, and Shores 2018a), who research the impact of opportunity on academic performance and achievement gaps in K-12 education and find that: (1) the variation in academic achievement among school districts is very large; (2) this variation is very highly correlated with the socioeconomic characteristics of families in the local community, with the strongest correlates of achievement gaps being local ethnic differences in parental income, local average parental education levels, and patterns of ethnic segregation; (3) the association between community socioeconomic status and academic performance grows steeper as children progress through school; (4) ethnic disparities in academic performance are large, both overall and within

[1] Unfortunately, candidate-level data on many relevant variables, e.g., number of science, math, and AP, courses offered and quality of counselling at high school, are not available. We try to capture these via measures of socioeconomic environment, community segregation, and university characteristics. Please see the methodology section for operational measures of these variables. We acknowledge that some prior research, e.g., Bline et al. (2016a) and Trinkle et al. (2016), has included select opportunity variables in their analyses.

individual school districts; and (5) part of the within-district ethnic achievement gaps are the result of ethnic disparities in family socioeconomic background. They conclude that socioeconomic and segregation context is a powerful force shaping children's educational opportunities and success.

Socioeconomic characteristics of families is also shown to affect performance in higher education. For example, Long and Mabel (2012) find that persistence in Ohio public four-year colleges is associated with family income level. Specifically, the bottom 20% of income distribution students have the lowest persistence rate to college degree completion. The persistence gap is significant between the top 20% and bottom 20% of income distribution even in cases where low-income students have scored 25 or higher on the ACT exam.

Socioeconomic characteristics of families, e.g., parental income and education levels, and patterns of ethnic segregation, greatly influence the type of k-12 school and college students attend, which in turn determine the opportunities available to them during and after college. For example, more selective and larger colleges provide their student with access to good counselling, relevant internships, employment by larger accounting firms who provide mentorship and financial support for CPA exam prep-courses. If African-American, Hispanic, and female candidates are unable to attend, or do not attend, these types of schools, then their performance on the CPA exam may suffer, partly or completely explaining the observed differential performance across ethnicity and gender.

As important as they are, the roles of socioeconomic factors, community segregation, and other measures of opportunity, in the CPA exam performance have not been addressed in the accounting literature. The current study fills this void. Using demographic and exam data from the National Association of State Boards of Accountancy (NASBA) over the period 2005-2016, university characteristics data from the Department of Education, and socioeconomic and segregation measures from Center for Educational Policy Analysis at Stanford University, univariate analysis reveals that a higher (lower) percentage of Caucasian candidates pass all four sections of the exam (drop after the first attempt or first section) than other candidates. Similarly, a higher (lower) percentage of males pass all four sections (drop after the first attempt or first section) than females. In addition, candidates from nonprofit, more selective, traditional (face-to-face), and AACSB accredited schools, perform better than those from for-profit, less-selective, online, and non-AACSB accredited schools, respectively. As for variables unique to this study, the results indicate that candidates that are from more affluent and less segregated communities and those that attend higher-tuition schools and schools with lower percentage of Pell Grant recipients, perform better than their counterparts. Importantly, a higher percentage of African-American, and to a lesser extent Hispanic and female, candidates come from a less affluent and more segregated communities and attend non-AACSB schools, for-profit schools, schools with lower average SAT scores, schools with higher percentage Pell Grant recipients, and schools with tuitions less than $20,000. Therefore, African-American, Hispanic, and female candidates enjoy less opportunity (e.g., less access to top teachers, mentorships and internship opportunities), which likely affects their performance.

We next develop three logistic regression and a survival models to test our hypothesis in a multivariate setting. The results of logit analyses indicate that candidates from more affluent and higher level of segregation of the Hispanic community who attend more selective and higher-tuition schools that offer graduate degrees have a greater (lower) probability of passing all four sections of the exam (dropping after the first attempt or section); and those from higher level of segregation of the African-American community who attend schools with higher proportion of Pell Grant recipients, larger, nonprofit, and AACSB-accredited have a lower (higher) probability of passing all four sections (dropping after the first

attempt or section). What is most notable is that controlling for opportunity variables, African-American and Hispanic candidates have a significantly higher (lower) probability of passing all four sections (dropping after first attempt or section) than Caucasians. Survival analysis results are similar to the logistic regression results, where the dependent variable is "Pass all four sections". In general, these results support our hypothesis that African-American and Hispanic candidates' underperformance on the CPA exam is related to socioeconomic conditions, community segregation, and other opportunity variables.

Our study contributes to the accounting and education literature in three main respects. First, it extends prior research in accounting by examining a more comprehensive set of factors that are potentially related to performance on the CPA exam. This reduces the likelihood of omitted correlated variables problem.[2] To the best of our knowledge, such a comprehensive study has not been performed. Equally important, by examining the impact of socioeconomic and other opportunity-related variables on CPA exam performance, our study extends the education literature. Specifically, it provides evidence on whether opportunity affects performance after graduation (on professional exams), rather than prior to graduation as shown by Reardon and others (e.g., Reardon 2016; Reardon et al. 2018a; Long and Mabel 2012). Third, the consistency of the results across both logistic regression and survival analyses (see the methodology section below) renders the study's conclusions about the role of opportunity in the CPA exam performance more credible. This is especially important because the non-parametric survival analysis avoids the problems resulting from the violation of the distributional assumptions of the ordinary least squares regression analysis.

Our findings have implications for the accounting profession, academia, and future research in this area. First, the accounting profession will benefit by understanding the myriad factors that relate to its ability to attract and retain a diverse talent pool in today's marketplace – an expressed top priority.[3] Specifically, the employers of future accounting talent can identify, and invest resources in, those programs that are effectively addressing opportunity gaps and expanding access to critical resources for successful performance on the CPA exam and long-term professional career success. Second, future regulations for the educational requirements to sit for the CPA exam, and to gain licensure, can consider potential barriers that exist for various population segments. Third, academic accounting programs of all types have a vested interest in assessing the impact that their curricular and co-curricular strategies have on the professional certification success of their students. Knowing that the achievement gaps on the CPA exam are mostly due to opportunity, educators, policy makers, professional firms, and CPA societies and State Boards, can focus on developing and offering educational and professional development strategies for all. This would allow professional firms and others to meet their diversity/inclusion goals, and to make our society prosperous overall. Finally, future research on the

[2] Omitted correlated variables result in biased coefficients (in sign and/or magnitude) in multivariate models. For example, if opportunity is related to performance and different ethnicities enjoy different opportunities, then excluding opportunity from the multivariate models causes the coefficient for ethnicity to incorporate the effect of opportunity, resulting in biased coefficient for ethnicity.

[3] As Jeanette Franzel, a former PCAOB member, stated "The ability to retain and develop top talent and achieve diversity in leadership ranks today is imperative to better positioning firms for the future. Firms that deal seriously with these issues also will have the advantage of increased access to larger talent pools. The complexity of auditing – and the vast responsibilities of firms in providing assurance over financial reporting for the benefit of investors and the markets – requires harnessing the talents and energies of a diverse workforce."

success factors in the CPA exam performance should control for the variables found to significantly impact performance.

Prior research and expected contributions

a. Accounting Literature

A number of studies in the accounting literature have partially addressed the issue of the CPA exam performance relative to ethnicity, gender, and university and candidate characteristics. For example, Booker (2003) examines participation and performance on the CPA exam by African-American candidates during 1997-1999. Descriptive statistics reveal that the majority of African-American candidates were female (between 57% and 66%), with median age of about 30, and 30% having earned 150 hours or more of college credit. The overall pass rates for first-time and repeat candidates ranged from 11% to 11.8% for the period of the study. These rates were below pass rates of other ethnic groups during this time period. Booker (2005) also reports that the number of African-American candidates increased between 2000 and 2002 with females comprising between 62% and 65% of this pool. Success factors cited for African-American candidates who passed the CPA exam during this period included internships, mentoring, scholarships, and co-op programs.

Enofe (2010) summarizes the literature on African-American experience with the CPA profession and concludes that economic, social, and cultural inequalities help explain the lack of this ethnic group in the accounting profession. However, there is no empirical analysis in this study to substantiate this conclusion. The current study attempts to empirically determine the role of socioeconomic factors in the CPA exam performance.

Brahmasrene and Whitten (2001) find a gender gap in the CPA exam performance among Indiana candidates in 1998. Univariate tests show that females were a larger percentage of candidates but a smaller percentage of successful candidates. The logistic regression results confirm that in that year males were more successful on the CPA exam than females.

Myers, Franklin, Lepak, and Graham (2018) provide a literature review of gender gap performance outcomes in several disciplines and extends this literature to consider CPA exam pass rates. Their study further substantiates earlier research findings of the significantly lower female exam pass rates for the period 2008-2015. The authors hypothesize that this result may be due to different information processing styles between genders, whereby males score higher with objective questions and females score higher with constructed response (e.g., simulation) questions. This finding calls for additional study following CPA exam emphasis changes that took place in 2017 to focus more on higher-order skill sets of candidates.

Several recent studies have examined the association between CPA exam performance and various characteristics of educational institutions candidates attended. Morgan (2015) finds that online accounting programs have much lower average CPA pass rates than their face-to-face counterparts with equivalent student selectivity at admission. Using t-tests assuming unequal variances, Bunker and Harris (2014) also find that candidates who attended predominantly online accounting programs perform significantly worse than candidates from both AACSB-accredited business schools and traditional, non-AACSB-accredited business schools on overall CPA exam pass rates and average CPA exam scores. Bunker, Cagle and Harris (2014) extend Bunker and Harris (2014) and use t-tests assuming unequal

variances to show that that CPA exam candidates who attended schools with separate AACSB-accredited accounting programs achieve higher overall pass rates and higher average exam scores than those from AACSB-accredited business schools that do not have separate accounting accreditation. The results of all three of these studies, however, are likely affected by confounds (other than student selectivity which Morgan controls for).

Using a large sample from NASBA of CPA exam candidate sittings from 2005-2013, Trinkle et al. (2016) develop a survival model using seven factors that had been identified in previous studies as significant determinants of CPA exam success. They find that candidates were more likely to succeed on the CPA exam if they were male, younger, received a degree from a college or university with an AACSB accredited business school and separately AACSB accredited accounting program, and received a degree from a private college or university. Their results support the gender gap phenomenon of earlier studies in that males were significantly more likely than females to pass the individual sections of the exam (except BEC), and 7% more likely to pass the complete exam.

Bline, et al. (2016a) develop a regression model to investigate the link between three accounting faculty characteristics and performance on the CPA exam, controlling for seven other variables including gender. The results based on nearly 700,000 first time exam sittings taken during the period 2005–2013, indicate that faculty expertise, research productivity, and CPA certification status are positively related to candidate exam score. They also indicate that candidates with a graduate degree score significantly higher on all four sections of the exam. The results regarding gender, age, and AACSB accreditation are consistent with those of Trinkle et al. (2016a); and that regarding selectivity, as proxied by the schools' average SAT score for incoming freshmen is consistent with those of Raghunandan, et al. (2003). Bline, et al. attribute the finding on age to older candidates being further removed from college coursework, or to other demographic characteristics of older candidates that may be negatively related to performance.

Finally, Mittelstaedt and Morris (2017) use NASBA data over 2005-2014 to develop three logit models, where the dependent variable is section score, section pass/fail, and pass all four sections, to examine CPA exam performance of for-profit vs. nonprofit university graduates. Controlling directly for year taken, exam type (first time/retake), exam section, gender, age, and accreditation status in the models, they find that graduates from nonprofit educational institutions outperform for-profit institutions' graduates on all four sections of the CPA exam. The coefficient of gender, age, and AACSB non-accreditation are negative in all three logit models, implying as in Trinkle et al. (2016) that younger, male, and AACSB-accredited school candidates perform better on the exam.

In addition to university characteristics, Bline, Perreault, and Zheng (2016b) link CPA exam performance to the order in which exam sections are attempted. They find that candidates who took the FAR section first successfully completed all exam parts sooner than those candidates who selected a different section order, over the period 2005-2013. They also find that females were less likely to take the FAR section first. Bline, et al. conjecture that the order in which candidates prefer to take the exam is related to their risk preferences. They go on to say that "a significant body of prior research indicates that males are typically more risk-seeking while females are typically more risk averse… To the extent that such risk preferences impact candidates' ordering of the exam (e.g., perhaps males are more likely to take the sections of the exam perceived as more challenging first) this may help to explain the significant performance gap between the genders that our [Bline, et al.] findings demonstrate".

No prior research has investigated the association between performance on the CPA exam and socioeconomic, community segregation, and other geography characteristics. Another variable that has not been investigated by prior research, but may be related to CPA exam performance, is the lapse of time between graduation and sitting for the exam. The longer a candidate waits to attempt the exam, the further removed is the candidate from college coursework, which the CPA exam largely focuses on. The delay in taking the exam is therefore likely to affect performance negatively.[4]

To sum, while many of these variables have been previously studied, they have generally been studied in isolation from others, using data for a short period and often data from only one state, and some are not empirically grounded. The current study extends prior studies on CPA exam performance by examining a larger set of potential explanatory variables including measures of socioeconomic status and community segregation, over an extended time period (2005-2016). It further employs various multivariate models to ensure robust results. The opportunity variables, especially socioeconomic conditions and segregation measures, are added because the education literature, briefly described in the next subsection, shows they affect performance in k-12, standardized exams such as SAT, and college. Although some of these effects may be less pronounced after college graduation in the select group that chooses to sit for the CPA exam, some residual effect are very likely to persist.

b. Education Literature

Education research has addressed exam performance relative to ethnicity, gender, socioeconomic conditions and segregation. Research on ethnicity shows that Caucasians perform better than African-Americans and Hispanics (Benner and Wang 2014; Clayton 2010; Reardon et al. 2018a). Gender research shows that females generally perform better in k-12 and university, while males generally perform better on standardized tests, including professional exams across racial groups (Fryer and Levitt 2009; Robinson and Lubienski 2011; Sohn 2012). However, studies of professions other than accounting with rigorous standards have found that, after controlling for predictive factors, gender is no longer a significant variable (see e.g., Stewart et al. 2006). Finally, a recent study by Reardon, Kalogrides, Fahle, Podolsky, and Zárate (2018b) finds that the average school district in their sample of 10,000 school districts has no gender achievement gap in math, but a gap of roughly 0.23 standard deviations in language arts that favors females. They further find that math gaps tend to favor males more in socioeconomically advantaged school districts and in districts with larger gender disparities in adult socioeconomic status.

Several studies find socioeconomic factors to be key predictors of children and adolescents' academic progress and ultimate educational success (e.g., Benner and Wang 2014; Lareau 2003; Long and Mabel 2012; Siegler, Duncan, Davis-Kean, Duckworth, Claessens, and Engel 2012; Kao and Thompson 2003).

[4] We do not consider the 150-hour requirement because (1) while nearly all states require the completion of 150 college credit hours to receive a license during the period of our study, many states allow candidates to take the CPA exam after attaining only a bachelor's degree, (2) candidates can sit for the exam in a state other than where they reside, and (3) prior research suggests that the implementation of the 150-credit hour requirement had little effect on the CPA exam pass rates (Allen and Woodland 2006; Trinkle, et al. 2016a). Our analysis instead considers the type of degree (graduate vs. undergraduate) and SAT scores. In our judgment, holding a graduate degree is a more relevant explanatory variable than the state education requirement. And, research shows that SAT scores, a measure of student aptitude, is a stronger indicator of exam success (Grant et al. 2002). Nevertheless, to test the robustness of our results, we control for the effect of varying state education requirement, by adding a state indicator variable (State_indicator) to our multivariate models. The results hold.

Recent research by Reardon, et al. (2018a) also reveals that the variations in performance in grades 3-8 is partially explained by socioeconomic and segregation factors. Crosnoe (2009), Kao and Thompson (2003), Lareau (2003), Lee (2002), Orfield and Lee (2004), and Siegler et al. (2012), further demonstrate that the conjunction between ethnicity and socioeconomic status is the most significant cause for the contrast in K-12 academic performance.

In Brown vs. Board of Education, the Supreme Court decided that "separate but equal" was, by definition, not equal. This case led to the desegregation of America's public schools and since then the promotion of ethnic and socioeconomic diversity in K-12 schools have been major goals of federal and state policy (Benner and Wang 2014). Unfortunately, "while facilities are now, in theory, integrated, the educational opportunities for students often are not" (Clayton, 2010, 3). The primary reason why these opportunities are different is that, for the most part, public students attend schools close to where they live due to personal reasons or lack of community resources. As a result, many young African-Americans attend ethnically segregated schools with more low-income students (Orfield and Lee 2004). In fact, Orfield and Yun (1999) show that there is a very high poverty rate among students in both segregated African-American and Hispanic schools. Thus, schools are effectively segregated by socioeconomic status and ethnicity.

Three other factors directly affect opportunity. These are tracking (grouping of students by their academic abilities), resource allocation, and faculty turnover (Clayton 2010). Academic tracking limits the opportunity for students to take higher level, honors, and AP classes, even those that attend schools that offer such courses. The second and the third factor are related. Inadequate financial resources limits the ability of schools to recruit and retain good teachers, to offer higher level courses, and to provide counseling and extracurricular services to their students.

Schools that enroll high proportions of ethnic minority and economically disadvantaged students have had lower financial resources because "No Child Left Behind" and "Race to the Top" programs, which had the goal of improving America's Public Education System, used standardized test scores as a basis for funding schools and districts. Because predominantly African-American and Hispanic schools serve students from disadvantaged neighborhoods, they both produce lower test scores and are inadequately funded. Yet, these are the same schools that need financial help the most. These schools have a challenging time recruiting and retaining teachers and suffer from high teacher-turnover rates. Inadequate funding has had a direct impact on student opportunity and achievement as classes and members of faculty and staff are cut.

The most recent educational policy initiative of the federal government, Common Core, seeks to improve teaching, learning, and equity by setting common goals for all students in a specific subject across a specific grade level throughout the country. However, Common Core also cannot ensure that opportunities are equal because again opportunities vary by zip code, and schools that are segregated by socioeconomic status and ethnicity have less opportunities (Kornhaber, Griffith, and Tyler 2014, 4).

Exacerbating the problem, students who attend more disadvantaged schools tend to perform worse on a variety of developmental indicators, including optimism, school engagement, achievement growth, and educational attainment," (Benner and Wang 2014, p. 1612). These four factors affect students' performance in k-12, and potentially in college, professional exams, and their general livelihoods.

To sum, education research shows that performance varies across the country and this variation is contributable mainly to ethnicity, gender, socioeconomic status, and segregation. This research further

shows that the conjunction between ethnicity and socioeconomic status, what is referred to as "double disadvantaged", is the most significant cause for achievement gaps.

Hypothesis

The accounting and education research described in the previous section indicates that ethnicity and gender gaps exist, that socioeconomic and segregation factors affect academic opportunity and academic performance at all level, and that the conjunction between ethnicity and socioeconomic status is the most significant cause of achievement gaps. CPA exam candidates face different opportunities throughout their education and afterward. These opportunities span across family characteristics, the community they grow up in, the k-12 schools and colleges they attend, and their professional work environment. Socioeconomic characteristics of families, e.g., parental income and education levels, and patterns of ethnic segregation, greatly influence the type of k-12 school and college students attend, which in turn determine the opportunities available to them during k-12, college, and after college. For example, more selective, traditional, AACSB accredited schools, and larger schools provide their students with access to good counselling, relevant curriculum and experienced faculty, relevant internships and coops, employment by larger accounting firms who provide mentorship and financial support for CPA exam prep-courses. If a larger percentage of African-American, Hispanic, female candidates grow up in less affluent, segregated communities; and attend less selective, online, non-AACSB-accredited, smaller, lower-tuition, and for-profit colleges, either due to lack of resources, preparedness (e.g., no AP courses and low SAT scores), or lack of mentors and counselors at home or at high school; then their performance on the CPA exam may suffer, partly or completely explaining the observed differential performance across ethnicity and gender. Therefore, we hypothesize that opportunity is the main driver of CPA exam performance.

Formally stated, we hypothesize that:

Candidates' performance on the CPA exam is related to their socioeconomic status, community segregation, and other measures of opportunity.

Sample, variables, and methodology

Table 1 shows the base sample, i.e., the number of exam sections[5] initially included in the sample by ethnicity and gender; those excluded for various reasons; and the final sample. The base/initial sample comprises exam sections over 2005-2016 taken by candidates with degrees from a US college or university and a score of at least 20 points. As expected, many exam sections were excluded from the sample due to data availability requirements. The most recent studies (Bline, et al. 2016a; Mittelstaedt and Morris 2017; Trickle, et al. 2016) have used sittings up to 2014 CPA exams. Other studies have used data from the 1990s and early 2000s. Given the changing population demographics and the documented inability of the profession to achieve its diversity/inclusion goals, any policy recommendation would benefit from more recent and a longer sample period horizon. Ideally, a continuous/dynamic research

[5] The CPA exam has four sections: Auditing & Attestation (AUD), Business Environment & Concepts (BEC), Financial Accounting & Reporting (FAR), and Regulation (REG). Students take the computer-based exam one section at a time, and after passing their first section, they have 18 months (548 days) to complete the remaining sections. After 18 months, they have to retake the previously passed sections outside the 18-months window.

program (as Reardon and his colleagues at Stanford and Harvard are doing in the k-12 domain) on the CPA exam results using this or an evolved methodology to detect patterns in performance and achievement gaps as a basis for policy recommendations and strategic action plans would be even more enlightening.[6]

[Insert Table 1]

The list of variables we consider in our analyses and their sources appear in Table 2. It includes socioeconomic and community segregation measures from the Center for Education Policy Analysis (CEPA) database at Stanford; university characteristics from the Department of Education (DOE) and NASBA databases, as well as other variables from NASBA that are theoretically or empirically shown to significantly influence candidates' performance on the CPA exam. About half of these variables have been considered before, but unfortunately in isolation from other correlated variables. Most studies have examined one or a few variables. We will examine these variables simultaneously as described below to overcome the problems with omitted correlated variables.

[Insert Table 2]

Our hypothesis proposes that candidates' performance on the CPA exam is related to socioeconomic status, segregation measures, and other measures of opportunity. To test our hypothesis, we first examine performance on CPA exam across socioeconomic index, community segregation, ethnicity, gender, and geography; and correlation between various variables. We then develop the following logistic regression and survival analysis models using as independent variables socioeconomic, segregation, other opportunity variables, as well as variables that prior research on CPA exam performance found to be significant:

$$
\begin{aligned}
Prob\,(&Pass\_all\_four\_sections, Drop\_after\_first\_attempt, Drop\_after\_first\_section) \\
&= \beta_0 + \beta_1 Ethnicity\_AA + \beta_2 Ethnicity\_HI + \beta_3 Gender + \beta_4 Age + \beta_5 Lapse \\
&+ \beta_6 Num\_attempts\_total + \beta_7 First\_section\_attempted + \beta_8 AACSB + \beta_9 Deg\_IRS \\
&+ \beta_{10} Deg\_high + \beta_{11} School\_high + \beta_{12} Enrollmen\_high + \beta_{13} Tuition\_high \\
&+ \beta_{14} SatVerbal75 + \beta_{15} Pellperc + \beta_{16} Sesall + \beta_{17} Hswhtblk + \beta_{18} Hswhthsp \\
&+ \beta_{19} Ethnicity\_AA * Sesblk + \beta_{20} Ethnicity\_HI * Seshsp + \beta_{21} Ethnicity\_AA \\
&* Hswhtblk + \beta_{22} Ethnicity\_HI * Hswhthsp + \beta_{23} Year\_scheduled
\end{aligned}
$$

Pass_all_four_sections, Drop_after_first_attempt, and Drop_after_first_section are measures of performance; variables on the right hand side are independent variables potentially affecting the CPA exam performance (starting row 4 in Table 2) and their interactions. Year_scheduled is the year fixed-effect.

[6] For example, beginning in 2017, the CPA exam underwent a number of changes. These changes emanated from a practice analysis launched by the AICPA in early 2014 on the knowledge and skills needed by newly licensed CPAs while performing tasks such as planning and reviewing the work of others. They include an increased focus on higher-order cognitive skills and professional skepticism such as analysis, evaluation, and application; an increased emphasis on task-based simulation exercises; and an increase in the length of both the BEC and REG sections of the exam. Undoubtedly, these changes will impact performance and achievement gaps and potentially policy recommendations and strategic action plans.

The logistic regression models developed using Equation (1) estimate the probability that a candidate will pass all four sections of the exam, drop after first attempt, or drop after first section as a function of socioeconomic, segregation, and opportunity variables, controlling for variables that prior research on CPA exam performance found to be significant. The survival analysis predicts the probability that a candidate will pass all four sections of the CPA exam in 18 months (548 days) given that the candidate has not yet passed all four sections of the CPA exam prior to the specific day. We estimate the survival function using both the Cox proportional hazard and the Kaplan-Meier methods. Cox proportion hazard model enables us to test the effect of other independent variables on survival times of different candidate groups, just like the multiple regression model. The Kaplan-Meier method of statistical treatment of survival times makes proper allowances for those observations that are censored; it also makes use of the information from these subjects up to the time when they are censored.

AACSB accredited schools are presumed to have higher quality programs and faculty and to attract better students and their graduates are more heavily recruited by large firms. As such, graduates of the AACSB schools are expected to perform better on the CPA exam. Findings of prior research support this assertion (see, e.g., Bline et al. 2016a; Morgan 2011; Self et al. 2013). Education provided by larger, higher-tuition universities with their generally better funded programs are expected to be superior to that of smaller universities, which tend to have less resources. Likewise students from more selective universities are more likely to pass the CPA exam because of the quality of their programs and faculty (Boone, Legoria, Seifert, and Stammerjohan 2006). Related to selectivity is the percentage of students who receive Pell Grants. Generally, less selective schools have higher percentage of Pell Grant recipients. The quality of education at for-profit universities in general is lower. For-profit students have lower graduation rates. Graduation rates within 6 years of enrollment for first-time, full-time bachelor's degree students at for-profit institutions has averaged 22% relative to 55% in public colleges and 65% in private nonprofit colleges (Lewin, 2010). Students from for-profit universities also are less likely to pass licensing exams in nursing, cosmetology, and law (Field, 2011) and in the CPA exam (Mittelstaedt and Morris 2017). Exam candidates that hold a graduate degree have the advantage of additional knowledge, but the disadvantage of being further away from undergraduate program coursework.

The four interaction terms are included because education research (e.g., Lareau 2003; Siegler et al. 2012; Kao and Thompson 2003; Lee 2002) shows a conjunction between socioeconomic status and ethnicity to be the most significant cause for the contrast in academic performance. Reardon (2016) also concludes that one reason for the large ethnic disparities in achievement may be that even among those who live in districts with similar socioeconomic conditions, African-American and Hispanic students are poorer, on average than Caucasians. Finally, year fixed effect is included to account for changes in the exam over time.

In addition to ethnicity, gender, socioeconomic status, community segregation, and other opportunity variables, the logistic and survival models above include four control variables: Age, Lapse, number of attempts, and first section attempted. Age could affect CPA exam performance in both directions. Older candidates may perform worse because they are more likely to have families or professional responsibilities than younger candidates, and are usually further away from graduation. However, older candidates are also generally more mature and have better study habits (Brahmasrene and Whitten 2001). Related to age is Lapse, the length of time between graduation and exam date. We would expect Lapse to have a negative effect on performance because the CPA exam is largely an undergraduate exam. The longer a candidate waits to attempt the exam, the further removed is the candidate from college coursework, and the lower the chances of success. Num_attempts_total is used as another control variable as candidates would be expected to become more familiar with the exam as they take it

more often and prepare for the next sitting. The inclusion of First_Section_Attempted is motivated by the results of Bline et al. (2016b), documenting that CPA exam performance is related to the order in which exam sections are attempted.

Results

a. Univariate analysis

Descriptive statistics (mean and standard deviation) for the sample are shown in Table 3. Descriptors are presented by ethnicity (first three sets of columns) and by gender (last two sets of columns). African-American and Hispanic candidates perform worse than Caucasians on the exam. For example, pass rates for all four sections for African-Americans, Hispanics, and Caucasians are 20, 32, and 47%, respectively. Further, relative to Caucasians, a higher percentage of African-American and Hispanic candidates drop after the first attempt (20 and 15 vs. 8%) and after the first section (5 and 4 vs. 2%).

African-American and Hispanic candidates are older on average than Caucasians (32.6 and 29.8 vs. 28.5 years) and they wait longer after graduation to sit for the exam (Lapse = 5.9 and 4.4 vs. 3.8 years). Relative to Caucasian candidates, on average a higher percentage of African-Americans attend non-AACSB schools (35 vs. 24%), for-profit schools (4.3 vs. 1.3%), schools with lower average SAT scores (507 vs. 547), schools with higher percentage Pell Grant recipients (37 vs. 28%), and schools with tuitions less than $20,000 (75 vs. 69%). For the most part, the type of school attended by Hispanic and Caucasian candidates are fairly similar. Interestingly, African-American females who take the exam comprise a larger percentage of all females than their male counterparts (7.9% vs. 5.5%). Similarly, Hispanic females who take the exam comprise a larger percentage of all females than their male counterparts (8.1% vs. 6.5%).

Differences on gender line are present but not as pronounced. For example, pass rates for all four sections for females and males are 41 and 47% and drop rates after first attempt are 10 and 8%, respectively. Female candidates are on average only slightly older than males (29.3 vs. 28.5) and take slightly longer time to sit for the exam (Lapse = 4.2 vs. 3.8 years). The type of school female candidates attend is also slightly different. Relative to males, on average a slightly higher percentage of females attend non-AACSB schools (27 vs. 24%), for-profit schools (2 vs. 1%), and schools with tuitions less than $20,000 (70 vs. 68%). These (univariate) results are consistent with the results of prior studies and provide comfort that the multivariate results in our study are not driven by differences in the sample.

[Insert Table 3]

The most interesting results relate to variables that have not been examined before. The average socioeconomic index (Ses_all) for African-Americans, Hispanics, and Caucasians are -0.15, 0.03, and 0.33, respectively. The socioeconomic index ranges from -4 to 4, with 0 being the national average socioeconomic condition. The values of the index indicate that African-American candidates are at a disadvantage compared to the National average while Hispanic candidates are just about average. The segregation index between Caucasians and African-Americans (Seg_AA) is 24, 20, and 13% for African-American, Hispanic, and Caucasian candidates, respectively; and that between Caucasians and Hispanics (Seg_HI) is 19, 16, and 11%, respectively. This indicates that African-American and Hispanic candidates come from more segregated communities than Caucasians. Again, differences across gender are present but not as pronounced. These statistics, and those on Pell Grant percentages and type of school

attended as discussed in the previous paragraph, support the idea that African-American candidates and, to a lesser extent, Hispanic and female candidates come from less affluent and more ethnically segregated communities than their counterparts.

We extend our univariate analysis by examining the correlation coefficients between pairs of select variables. Almost all correlations shown in Table 4 have the expected signs. For example, consistent with the descriptive statistics and prior research, Ethnicity_AA and Ethnicity_HI are significantly and negatively related to Pass_all_four_sections (-0.13 and -0.07) and positively related to Drop_after_first_attempt (0.10 and 0.06), while Ethnicity_CA is positively related to Pass_all_four_sections (0.15) and negatively to Drop_after_first_attempt (-0.11). Gender, Age, and Lapse are significantly and negatively associated with Pass_all_four_sections (-0.06, -0.22, and -0.14); Gender and Age are positively associated with Drop_after_first_attempt, but Lapse is negatively associated with Drop_after_first_attempt (the coefficient between Lapse and Drop_after_first_attempt is -0.04). AACSB accreditation and nonprofit status (Deg_IRS) are positively related to Pass_all_four_sections, and negatively to Drop_after_first_attempt. As for variables unique to our study, Pellperc is negatively related to Pass_all_four_sections and negatively to Drop_after_first_attempt (-0.17 and 0.11); it is positively and significantly related to Ethnicity_AA and Ethnicity_HI (0.15 and 0.13). The Pellperc correlations suggest that African-American and Hispanic candidates attend schools with higher percentage of students on Pell Grants, the graduates of which perform worse on the CPA exam. Community segregation measures, Seg_AA and Seg_HI, are positively and significantly related to Ethnicity_AA and Ethnicity_HI (0.16 and 0.11), indicating that African-American and Hispanic candidates on average come from more ethnically segregated high schools. The correlation coefficients on university characteristics and community segregation together confirm that African-American and Hispanic candidates are from less affluent and more segregated communities.

[Insert Table 4]

b. Multivariate Analysis

The evidence presented in Tables 3-4 provides support for the idea that socioeconomic status, community segregation, and other opportunity variables affect performance on the CPA exam. We next test our formal hypothesis using the multivariate models shown as Equation (1). Specifically, in logistic regressions we model the probability that a candidate will pass all four sections of the exam, or drop after the first attempt or first section, as a function of various socioeconomic, segregation, and opportunity variables, controlling for variables that prior research on CPA exam performance found to be significant. In survival analysis, we model the probability that a candidate will pass all four sections of the CPA exam in 18 months (548 days) given that the candidate has not yet passed all four sections of the exam prior to a specific day.

The results for logistic regressions are reported in Table 5. The coefficients on most variables in Models 1-3 are significantly different from zero at the 1% level. In model 1, the only insignificant coefficients are those of the two interaction variables between socioeconomic index and ethnicity. This implies that the socioeconomic status of African-American and Hispanic candidates relative to Caucasians does not affect the performance of these candidates. Candidates from communities that are more affluent (Ses_all) and where segregation between Caucasians and Hispanics are higher (Seg_HI) who attend more selective and higher-tuition schools that offer graduate degrees (SatVerbal75 , Tuition_High, School_High), have a greater probability of passing all four sections of the exam; and those from

communities with higher level of segregation between Caucasians and African-Americans (Seg_AA) who attend schools with higher proportion of Pell Grant recipients (Pellperc), larger (Enrollment_High), nonprofit (Deg_IRS), and AACSB-accredited (AACSB) have a lower probability of passing all four sections. The signs of Deg_IRS and AACSB are contrary to our expectations and the findings of prior research, and may be due to multicollinearity (e.g., larger schools are normally AACSB accredited and offer graduate degrees).[7] What is most striking is that the coefficients for Ethnicity_AA and Ethnicity_HI are positive and significant. It appears that controlling for opportunity variables, African-American and Hispanic candidates have a significantly higher probability of passing all four sections than Caucasians. However, the coefficient for Gender is still negative, indicating the females are less likely to pass all four sections than males. In general, these results support our hypothesis that candidates' performance on the CPA exam is related to socioeconomic status, community segregation, and other measures of opportunity.

[Insert Table 5]

In Models 2 and 3, the coefficients for First_section_attempted other than BEC, nonprofit status of the university (Deg_IRS), and socioeconomic index (Ses_all) are not significantly different from zero, indicating that these variables are not related to the probability that a candidate will drop after the first attempt or first section. In addition, the segregation index between Caucasians and Hispanics (Seg_HI) is also not significant in Model 3. Focusing on significant results and opportunity-related variables, as expected, the signs of the coefficients for all variables in models 2 and 3 are opposite to their signs in model 1 with a few exceptions, mostly non-significant at the conventional levels. Lapse in model 2, nonprofit status of the school (Deg_IRS) and SatVerbal75 in model 3, and Ses_all in both models 2 and 3, are associated with a lower probability that a candidate drops after the first attempt or section. This partially explains the difference between our results in model 1 that candidates that attend nonprofit schools are less likely to pass all four sections and prior research finding to the contrary. Specifically, the results in prior research (e.g., Mittelstaedt and Morris 2017) may be driven by candidates who dropped after attempting the first section. Again, what is most striking is that the coefficients for Ethnicity_AA and Ethnicity_HI are negative and significant. That is, controlling for opportunity variables, African-American and Hispanic candidates have a much lower probability of dropping after the first attempt or section than Caucasians. However, the coefficient for Gender is positive, indicating the females are more likely to drop after the first attempt or section. These results further support those in model 1 in that candidates' performance on the CPA exam is related to socioeconomic status, community segregation, and other measures of opportunity.

Table 6 presents the odds ratios and probabilities (p) for each variable in Table 5 logistic regressions. An odds ratio greater than 1.0 for a dichotomous predictor variable indicates that the odds that a candidate associated with a value of 1 for the predictor variable will pass all four sections, drop after the first attempt, or drop after the first section (event odds), is greater than the non-event odds; with an odds ratio less than 1.0 indicating the opposite. The odds ratio for Ethnicity_AA in Model 1 is 2.51, suggesting that the odds of passing all four sections of the exam for an African-American candidate is 151% higher than the odds for other candidates. The odds ratio for Ethnicity_HI is 1.47, suggesting that the odds of passing all four sections of the exam for a Hispanic candidate is 47% higher than the odds for other candidates. Females have 22% lower odds of passing all four sections than males. In Models 2 and 3, the

[7] Although based on data from 1998-1999, Boone et al. (2006) find that accreditation is not significantly related to pass rate. They conjecture that the correlation between AACSB accreditation and pass rate is explained by student selectivity, a variable which is also included in our models.

odds ratio for Ethnicity_AA (Ethnicity_HI) are 0.42 and 0.84 (0.58 and 0.73), indicating that the odds that an African-American (a Hispanic) candidate will drop after the first attempt or first section is 58% and 16% (42% and 27%) lower than the odds for other candidates. Females have 13% and 8% higher odds of dropping after the first attempt or section than males.

[Insert Table 6]

For a continuous predictor variable, an odds ratio greater than 1.0 implies an increase in the event odds as a result of a one-unit increase in the predictor value, holding all other variables at a certain value. The odds ratio for Ses_all in Model 1 is 1.03, indicating that for a one-unit increase in Ses_all, we can expect a 3% increase in the odds of passing all four sections, holding all other variables constant. The odds ratio for Seg_AA, segregation between Caucasians and African-Americans, in Model 1 is 0.47, indicating that for a one-unit increase in Seg_AA we can expect a 53% decrease in the odds of passing all four sections, holding all other variables constant. In models 2 and 3, the odds ratio for Ses_all are 1.01 and 1.03, indicating that for a one-unit increase in Ses_all we can expect a 1% and 3% increase in the odds of dropping after the first attempt or first section. The odds ratio for Seg_AA (Seg_HI) are 1.88 and 2.22 (0.44 and 0.58), indicating that for a one-unit increase in Ses_all (Seg_AA) we can expect an 88% and 122% increase (56% and 42% decrease) in the odds of dropping after the first attempt or first section. The differences between the impact of community segregation on performance of African-American and Hispanic candidates are interesting in that segregation appears to help Hispanic candidates and not African-Americans.

The odds ratios for the interaction terms involving the socioeconomic status of African-American and Hispanic candidates relative to the overall community (in the zip code) in Models 1-3 are 1 or very close to it. This implies that the differences in the socioeconomic status of African-American and Hispanic candidates from the community do not affect their chance of passing all four sections or dropping after the first attempt/section. Community segregation on the other hand does impact the performance of African-American and Hispanic candidates relative to others. For example, in Model 1 the odds ratios for the interaction terms Ethnicity_AA * Seg_AA and Ethnicity_HI * Seg_HI reveal that a 1% increase in segregation (Seg_AA and Seg_HI) reduces the chances of passing all four sections of the CPA exam in 18 months by 36% and 56% for African-Americans and Hispanics, respectively.

Overall, the results of logistic regressions support the idea that poorer performance of African-American and Hispanic candidates' performance on the CPA exam is related to socioeconomic status, community segregation, and other measures of opportunity.

Survival analysis results are presented in Table 7. In the first three columns, we report the parameter estimates and the related p-values, as well as the hazard ratios minus one, for the Cox proportional hazards (parametric) model. The results largely support the logistic regression results. However, two opportunity variables that are significant in Model 1 logistic regression, namely, socioeconomic index and nonprofit status of the university (Ses_all and Deg_IRS), are insignificant. The interaction of Ethnicity and socioeconomic status of African-Americans relative to the overall community (Ethnicity_AA * Ses_AA) becomes negative and significant at 6% level. Additionally, the coefficient of number of attempts across all sections of the CPA exam (Num_attempts_total) changes from positive to negative. This sign change is expected as, the negative coefficient in the survival analysis indicates that a candidate who attempts the CPA exam more times in total is less likely to pass all sections the next time given that she or he has not already passed all four sections. A positive sign in the logistic regression indicates that a candidate who attempts the CPA exam more times is more likely to pass all four sections

of the exam. Both interpretations make logical sense and thus the difference in the sign of Num_attempts_total seems appropriate. The signs for two other variables, the nonprofit status of the university (Deg_IRS) and the interaction between ethnicity and socioeconomic index for Hispanic candidates (Ethnicity_HI * Ses_HI), also change; but these variables are insignificant in the survival model.

[Insert Table 7]

A hazard ratio is interpreted in much the same way as the odds ratio. The opportunity variables related to the type of school attended in general impacts a candidate's chance of passing all four sections of the CPA exam. A candidate who has not yet passed the exam has a 7.2%, 14.6%, and 1.1% lower chance of passing all four sections of the CPA exam on the next attempt if the candidate attended an AACSB accredited school, a larger school (Enrollment_High), and a school with a high proportion of Pell Grant recipients (Pellperc), respectively. A candidate's chance of passing the CPA exam on the next attempt is increased by 15.2% if the candidate attended a high tuition school (Tuition_High). The school's non-profit status (Deg_IRS) and selectivity (SatVerbal75) do not have a significant effect on the candidate's chance of passing all four sections in 18 months (548 days), although the coefficient of the latter is statistically significant.

The coefficients relating to socioeconomic status are not significant. The coefficient of socioeconomic status (Ses_all) is not statistically significant. Those of the interaction terms between Ethnicity and socioeconomic index for Hispanics (Ethnicity_HI * Ses_HI) and African-Americans (Ethnicity_AA * Ses_AA) are also not statistically or practically significant at the conventional levels, indicating that differences in the socioeconomic status of African-American and Hispanic candidates from the community do not affect their chance of passing all four sections on their next attempt. However, community segregation has a significant effect on candidates' chances of passing all four sections of the CPA exam. Specifically, a candidate associated with a 1% higher segregation index between Caucasians and African-Americans (Seg_AA) is 27% less likely to pass the CPA exam on the next attempt. This reduction in the chance of passing the CPA exam is increased further by another 28% for African-American candidates (Ethnicity_AA * Seg_AA). On the other hand, a 1% higher segregation of the Hispanics (Seg_HI) increases the probability that a candidate passes all four sections of the exam by 38%. However, the chances of Hispanic candidates passing the exam in those same communities (Ethnicity_HI * Seg_HI) decreases by 41%. This means that Hispanic candidates have a net 3% lower chance of passing all four sections as community segregation increases by 1%.

Consistent with the logistic regression results, we find that, holding all other effects constant, African-Americans and Hispanics are more likely to pass all four sections of the CPA exam than Caucasians. Specifically, African-Americans have 67% and Hispanics 24% higher chance of passing all four sections of the CPA exam on their next attempt than Caucasians. These results supports the hypothesis that differential candidates' performance on the CPA exam is related to socioeconomic status, segregation measures, and other measures of opportunity. However, as in the logistic regressions, the Cox proportional hazard model suggests that females have a lower probability of passing all four sections of the CPA exam.

The Cox Proportional Hazard model requires that the slopes of the hazard function are parallel. That assumption is relaxed with the non-parametric survival function. The last two columns of Table 7 report the Wilcoxon Chi-squared test statistics and the related p-values for the nonparametric survival function, which is estimated by the Kaplan-Meier method. The only variable that is not statistically

significant is the interaction term between ethnicity and the socioeconomic status for Hispanics (Ethnicity_HI * Ses_HI). The nonprofit status of the university (Deg_IRS), the socioeconomic status (Ses_all), and the interaction term between ethnicity and the socioeconomic status for African-Americans (Ethnicity_AA *Ses_AA), however, become significant in the nonparametric survival analysis.

The survival curves by Ethnicity are presented in Figure 1. The slopes of the survival functions for each ethnicity differs with Caucasians being the steepest and African-Americans being the flattest. They show that Caucasians have the lowest probability of passing all four sections of the CPA exam followed by Hispanics given they have not yet successfully passed the CPA exam for all days. This indicates the probability of successfully passing all four sections of the CPA exam as the interval between the first exam section passed and the next attempt increases, falls by less for African-Americans followed by Hispanics, with Caucasians having the more significant reductions in the probability of passing all four sections of the CPA exam over time.

[Insert Figure 1]

Figure 2 shows the survival functions for passing all four sections of the exam stratified by gender. The slope of the survival function for females is flatter than that for males, indicating that, compared to females, males experience larger reductions in the chance of passing all four sections of the CPA exam as time since passing the first section increases. This provides support for the hypothesis in that females have a higher probability of passing the CPA exam once socioeconomic, segregation, and other opportunity factors are taken into account. Given that that slopes of the survival functions do not appear to be parallel, the non-parametric results deserve more consideration.

[Insert Figure 2]

Like most empirical research, our study is subject to limitations. First, our measures of ethnicity, gender, and some other demographic data are self-reported and may be inaccurate. Second, due to unavailability of data, we had to use school or zip-code level, and not candidate level, data for some variables (e.g., SatVerbal75 and Pellperc). The associated measurement problems may have affected the results. Third, it is possible that a different set of opportunity variables potentially effect performance of female candidates. For example, greater family obligations of female candidates could potentially affect their performance. Finally, we were unable to obtain information about many opportunity variables, e.g., the availability of AP courses and counseling at high school level, availability of guidance and mentors at college, and access to a CPA exam review course prior to sitting for the exam. Therefore, the study is still subject to omitted explanatory variable, albeit at a much lower level than previous studies. While we see no reason to expect that the variables in our models are correlated with availability of AP courses and counseling at high school level, availability of guidance and mentors at college, etc.;[8] and believe that socioeconomic index and segregation measures capture the effects of missing information; future research may wish to explore how these specific factors impact CPA exam performance.

[8] An exception to this statement is the availability of CPA exam review course since accounting firms that provide free access to CPA review courses to their new hires normally hire a greater number at larger, AACSB schools.

Summary and Policy Recommendations

Professional accounting firms are excellent places for college graduates to launch their business careers. This type of environment allows for individuals to partake in life-long learning and thus reap the reward of multiple career opportunities during their lifetime. The key that opens the door to these opportunities is successful completion of the CPA exam, leading to licensure as a CPA. Employers are focused on hiring college graduates who have a plan on getting to and passing the CPA exam within the first two years of employment, with an eye on their diversity/inclusion goals. In fact, some firms even offer a generous bonus to employees who pass the exam within the first two years of employment. Given the demand for, and private returns to, CPA attainment, a key policy question is how to improve candidates' performance on the exam.

This study uses data on candidate-level demographic and exam results from the National Association of State Boards of Accountancy over the period 2005-2016, university characteristics from the U.S. Department of Education, and socioeconomic and community segregation measures from the Center for Education Policy Analysis at Stanford University, to examine the relation between opportunity, ethnicity, and gender and the CPA exam performance. In our multivariate analyses, we control for several factors that prior research has found to be significantly associated with CPA exam performance. As such, our analyses are based on a unique, and more comprehensive, set of explanatory variables for examining performance on the CPA exam.

Univariate analysis results are consistent with prior studies. They indicate that Caucasian candidates perform better than African-American and Hispanic candidates and male candidates perform better than females. Specifically, a higher (lower) percentage of Caucasian candidates pass all four sections of the exam (drop after the first attempt or first section) than other candidates. Similarly, a higher (lower) percentage of males pass all four sections (drop after the first attempt or first section) than females. In addition, candidates from nonprofit, more selective, traditional (face-to-face), and AACSB accredited schools, perform better than those from for-profit, less-selective, online, and non-AACSB accredited schools, respectively. As for variables unique to this study, the results indicate that candidates that are from more affluent and less segregated communities and those that attend more expensive schools and schools with lower percentage of Pell Grant recipients, perform better than their counterparts. Importantly, a higher percentage of African-American, and to a lesser extent Hispanic and female, candidates come from less affluent and more segregated communities and attend non-AACSB schools, for-profit schools, schools with lower average SAT scores, schools with higher percentage Pell Grant recipients, and schools with tuitions less than $20,000. Therefore, African-American, Hispanic, and female candidates enjoy less opportunity, which likely affects their performance.

Given these findings, it is important that conclusions and policy recommendations be based on multivariate models that control for the effect of opportunity on performance. To that end, we develop several logistic regression and survival models. The results of logit analyses indicate that candidates from more affluent and more segregation of the Hispanic communities who attend more selective and higher-tuition schools that offer graduate degrees have a greater (lower) probability of passing all four sections of the exam (dropping after the first attempt or section); and those from more segregation of the African-American communities who attend schools with higher proportion of Pell Grant recipients, larger, nonprofit, and AACSB-accredited have a lower (higher) probability of passing all four sections (dropping after the first attempt or section). What is most notable is that controlling for opportunity variables, African-American and Hispanic candidates have a significantly higher (lower) probability of passing all four sections (dropping after first attempt or section) than Caucasians. In general, these

results show that African-American and Hispanic candidates' underperformance on the CPA exam is related to socioeconomic conditions, community segregation, and other opportunity variables.

Parametric survival analysis based on Cox Proportional Hazard model indicates that African-American and Hispanic candidates have a much higher chance of passing all four sections of the CPA exam on their next attempt than Caucasians. It further indicates that females have a lower probability of passing the CPA exam on their next attempt than males. These results are consistent with the logistic regression models.

Non-parametric survival model yields the same results as Logistic regression and parametric survival models with respect to ethnicity, namely that African-American and Hispanic candidates have a much higher chance of passing all four sections of the CPA exam on their next attempt than Caucasians. However, the non-parametric survival graphs reveal that females are more likely to pass all sections of the CPA exam on their next attempt than males. Because the slopes of the survival functions are not parallel, the non-parametric results deserve more weight.

The results of this study have significant policy implications in a macro sense. Educational literature shows that socioeconomic status and community segregation could be significant in explaining CPA exam performance differences across ethnicity and gender. The empirical results of this study using a unique combination of data sets support these findings.

The business and moral cases for a highly qualified and diverse accounting talent pool, both in public and private settings, have been well publicized for years. The various constituencies that are invested in CPA exam performance outcomes all have vital roles to play in implementing these findings. To help meet diversity goals in the accounting profession, broadly defined, successful completion of the CPA exam (among other candidate characteristics) is an important prerequisite. First, the major institutional CPA constituencies such as NASBA and the AICPA, and accounting and financial service firms, should continue to lobby for broad macro policies that address their diverse talent needs. Specifically, to assist the various population segments within the profession that are the focus of this research these groups should articulate the impact of socioeconomic factors (e.g., median income, higher education attainment, poverty rates, single mother head of households, unemployment rates) on future diverse talent pools. Likewise, ethnic diversity of schools and neighborhoods should be considered in planning communities and school districts. Clearly these types of factors are long-term, general society issues but the accounting profession should be advocating for their importance relative to the need for a diverse future talent pool.

On a shorter term horizon, the CPA constituents (NASBA, State Accounting Societies, the AICPA, and accounting and financial service firms) should adopt new, or expand existing, intervention programs for identifying and nurturing diverse talent who have the limited opportunity factors cited in this study. Such programs should include career information sessions in high schools in conjunction with existing student advising centers. The importance of providing high school students with professionals as role models should be considered. These role models could share their own experiences and offer students advice on such matters as college selection, courses of study for accounting, and externships and internships, for example. This activity could have a significant impact on diverse students' career decisions towards accounting careers. These same efforts could be introduced, or expanded, on college campuses with significant diverse populations to engage students about accounting careers early in their university studies. Of course, the importance of the CPA exam to their eventual professional certification process would be included in this information flow to students. Indeed, the Pathways

Commission recommendation number five provides a suggestion to improve the ability to attract high-potential, diverse entrants into the accounting profession. An opportunity also exists for these organizations to provide students with data to inform them of CPA exam outcomes of accounting programs across the United States. Pell Grant recipients especially would benefit from the implementation of this recommendation. Further, the CPA constituents should support individuals seeking entry into the accounting profession by providing mentors who can advise and provide evidence of career advancement, summer programs for college students to expand their professional exposure, and scholarships to help with finances. Indeed these factors have often been cited in previous accounting and education literature as important.

Expanding the use of need-based financial aid to qualified, diverse students pursuing accounting would be a valuable supplemental resource to existing university financial aid pools. Such a private source of financial aid could be an important incentive to these students as an alternative to accumulating a larger student loan burden during undergraduate college studies. Alternatively, this type of financial aid could be directed specifically to completing the additional thirty credits needed for CPA licensure and/or CPA exam review courses.

These efforts may be especially needed at those universities where there is little or no exposure to the accounting profession and its professionals which is a great advantage to students at schools where this presence is typical. Schools lacking this professional exposure are likely to have a larger diverse population segment. Students who do not attend more privileged universities need an advocate to shepherd them through the career placement and certifying exam processes. To the extent that budgets are limited for professional and private firms to undertake this level of investment, perhaps creating an advocacy organization similar to Sponsors for Educational Opportunities (SEO) would help diverse students to "land on their feet" with respect to the CPA exam and the accounting profession. Such an organization could provide the additional thirty credits for CPA licensure by integrating CPA exam content, as well as new competencies and "soft" skills into contextualized courses to remediate the deficiencies related to socioeconomic and segregation factors.

A recent study by Castleman, Long and Mabel (2018) offers evidence that once students enroll in college, availability of need-based financial aid in STEM increases STEM credit completion by 20 to 35 percent among academically-ready students. Could the same happen for CPA attainment if need-based financial aid were available for the additional 30 credits needed for CPA licensure and CPA exam review course?

We recognize that the accounting profession has already initiated several programs to assist exam candidates. For example, at http://www.aicpa.org/nextcpaexam, candidates can take sample tests as well as see the new Document Review Simulation (DRS). At http://www.ThisWayToCPA.com/examinfo, candidates can find tools to learn the requirements for the exam and licensure process, get advice from CPAs who successfully passed the exam, and start planning their study schedule. These programs are valuable in addressing the information and opportunity factors that are the subject of this study and we strongly support their continuation.

The education community also has an important role in addressing the diverse talent pool goal for the accounting profession. Obvious responsibilities include relevant curricula content and skill development to assist students in becoming successful CPA exam candidates. Beyond that, however, universities and their faculty can assist in promoting student exposure to professional accountants and professional accounting environments in general. Faculty can be an important information conduit about the CPA

exam and professional issues to all students, as well as explaining the marketability of the accounting skill set. These activities are certainly consistent with the socioeconomic and segregation factors reported in the current study.

References cited

American Institute of Certified Public Accountants (AICPA). 2017. *Trends in the Supply of Accounting Graduates and the Demand for Public Accounting Recruits*. Ewing, NJ: AICPA.

Benner, A. and Y. Wang. 2014. Demographic marginalization, social integration, and adolescents' educational success, *Journal of Youth Adolescence* 43: 1611–1627.

Bline D., S. Perreault, and X. Zheng. 2016a. Do accounting faculty characteristics impact CPA exam performance? An investigation of nearly 700,000 examinations, *Issues in Accounting Education* 31 (3): 291–300.

Bline D., S. Perreault, and X. Zheng. 2016b. An exploratory examination of order effects on CPA exam passage timeliness. *Journal of Accounting Education* 36: 65–74.

Booker, Q. 2003. African-American candidates for the CPA exam: Participation and performance. The CPA Journal 197 (8): 55-57.

Booker, Q. 2005. African American students and the CPA exam. *Journal of Accountancy* 199 (5): 60–64.

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

Brahmasrene, T. and D. Whitten. 2001. Assessing success on the uniform CPA exam: A Logit approach, *Journal of Education for Business* 77 (1): 45-50.

Bunker, R. and D. Harris. 2014. Online accounting degrees: An empirical investigation of CPA exam success rates. *Journal of Business and Accounting* 7 (1): 86-93.

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

Castleman, B., B. Long, and Z. Mabel. 2018. Can financial aid help to address the growing need for stem education? The effects of need-based grants on the completion of science, technology, engineering, and math courses and degrees. *Journal of Policy Analysis and Management* 37 (1): 136–166.

Clayton, J. 2010. Changing diversity in U.S. schools: The impact on elementary student performance and achievement, Education and Urban Society 43 (6): 671-695.

Crosnoe, R. (2009). Low-income students and the socioeconomic composition of public high schools. *American Sociological Review* 74: 709–730.

Enofe, A. 2010. Certified public accounting profession: The African-American experience. *International Journal of Business, Accounting, and Finance* 4 (1): 157-168.

Fahle, E. and S. Reardon. Forthcoming. How much do test scores vary among school districts? New estimates using population data, 2009-2015. *Education Researcher*.

Fryer, R. Jr. and S. Levitt. 2009. An empirical analysis of the gender gap in mathematics. American Economic Journal *Applied Economics* 2(2): 210-240.

Kao, G. and J. S. Thompson. 2003. Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology* 29: 417–442.

Lareau, A. 2003. *Unequal childhoods: Class, race, and family life*. Berkeley, CA: University of California Press.

Lee, J. 2002. Racial and ethnic achievement gap trends: Reversing the progress toward equity? Educational Researcher 31: 3–12.

Long, B. and Z. Mabel. 2012. Analysis of Ohio public four-year colleges. *Working Paper*, Harvard University.

Mittelstaedt, H and M. Morris. 2017. Academic achievement by graduates from for-profit and nonprofit institutions: Evidence from CPA exam performance. *Journal of Education for Business* 92 (4): 161-172.

Morgan, J. 2015. Online versus face-to-face: A comparison of CPA exam outcomes across matched institutions, *Journal of Education for Business* 90 (8): 420-426.

Myers, J., M. Franklin, G. Lepak, and J. Graham. 2018. The impact of gender and cognitive information processing models on CPA exam pass rates: A call for research. *Journal of Business and Educational Leadership* 7 (1): 59-71.

Orfield, G. and C. Lee. 2004. *Brown at 50: King's dream or Plessy's nightmare?* Cambridge, MA: Civil Rights Project of Harvard University.

Orfield, G. and J. Yun. 1999. *Resegregation in American schools*. Cambridge, MA: Civil Rights Project of Harvard University.

Raghunandan, K. Read, and C. Brown. 2003. The 150-hour rule: Does it improve CPA exam performance? *Managerial Auditing Journal* 18 (1): 31-38.

Reardon, S. 2016. School District Socioeconomic Status, Race, and Academic Achievement. *Working Paper*, Stanford University.

Reardon, S, D. Kalogrides, and K. Shores. 2018a. The geography of racial/ethnic test score gaps. *American Journal of Sociology*

Reardon, S., D. Kalogrides, E. Fahle, A. Podolsky, and R. Zárate. 2018b. The relationship between test item format and gender achievement gaps on math and ELA tests in 4th and 8th grade. *Education Researcher*

Robinson, J. and S. Lubienski. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal* 48(2): 268-302.

Siegler, R., G. Duncan, P. Davis-Kean, K. Duckworth, A. Claessens, and M. Engel, M. 2012. Early predictors of high school mathematics achievement. *Psychological Science* 23: 691–697.

Sohn, K. 2012. A new insight into the gender gap in math. *Bulletin of Economic Research* 64(1): 135-155.

Stewart, E. 2006. Family- and individual-level predictors of academic success for African American students: A longitudinal path analysis utilizing national data. *Journal of Black Studies* 36 (4): 597-621.

Trinkle, B., J. Scheiner, A. Baldwin, and G. Krull. 2016. Gender and other determinants of CPA exam success: A survival analysis. *The Accounting Educators' Journal* 26: 101-118.

Table 1
Sample

This table reports the base sample in this study, which consists of all CPA exam sections attempted by candidates having their higher degree from a US college or university and scoring at least 20 points by ethnicity and gender over 2005-2016; those excluded for various reasons; and the final sample size.

	Caucasians	African-Americans	Hispanics	Male	Female
Base sample – total NASBA population of CPA exam sections from 2005 to 2016 with the higher degree from a US college or university and a score of >= 20	752,036	56,818	73,928	429,525	453,257
Less: Exam sections retaken	277,596	27,157	31,960	153,591	183,122
Less: Exam sections with invalid or missing required NASBA data	50,602	2,954	10,195	31,252	32,499
Less: Exam sections with invalid or missing required DOE data	33,306	1,958	2,550	20,217	17,597
Less: Exam sections with invalid or missing required CEPA data	14,929	680	554	7,238	8,925
Final sample – exam sections	375,603	24,069	28,669	217,227	211,114
Final sample – number of candidates	118,250	9,169	10,076	69000	68,495

Table 2
Variable definition

Variable label	Variable description	Source*
Pass_all_four_sections	an indicator variable that = 1 if the candidate passed all four sections of the CPA exam and zero otherwise	NASBA
Drop_after_first_attempt	an indicator variable that = 1 if the candidate dropped after the first attempt and zero otherwise	NASBA
Drop_after_first_section	an indicator variable that = 1 if the candidate dropped after attempting the first section and zero otherwise	NASBA
Ethnicity_AA	an indicator variable that = 1 if the candidate is African-American and zero otherwise	NASBA
Ethnicity_HI	an indicator variable that = 1 if the candidate is Hispanic and zero otherwise	NASBA
Gender	an indicator variable that = 1 if the candidate is female and zero otherwise	NASBA
Age	age of candidate at the time of last exam	NASBA
Lapse	number of years between graduation and last exam date	NASBA
Num_attempts_total	total number of attempts across all sections	NASBA
First_section_attempted	a categorical variable that equals the exam section the candidate first attempted, AUD, BEC, FAR, or REG	NASBA
AACSB	an indicator variable that = 1 if the last school attended is accredited by the AACSB and zero otherwise	NASBA
Deg_IRS	an indicator variable that = 1 if the last school attended is nonprofit and zero otherwise	NASBA
Deg_High	an indicator variable that = 1 if the candidate's highest degree is at post-graduate level and zero otherwise	NASBA
School_High	an indicator variable that = 1 if the last school attended offers graduate programs and zero otherwise	NASBA
Enrollment_High	an indicator variable that = 1 if the enrollment at the last school attended exceeds 20,000	NASBA
Tuition_High	an indicator variable that = 1 if annual tuition at the last school attended exceeds $20,000	NASBA
SatVerbal75	75 percentile value of average SAT verbal score for the last school the candidate attended	DOE
Pellperc	percentage of students receiving Pell Grants at the last school the candidate attended	DOE
Ses_all	socioeconomic index for all in a given zip code and is the first principal component factor scores of the following six measures: median income, percent with a bachelor's degree or higher, poverty rate, SNAP rate, single mother headed household rate, and unemployment rate	CEPA
Ses_AA	socioeconomic index for African-Americans relative to socioeconomic index for all in a zip code	CEPA
Ses_HI	socioeconomic index for Hispanics relative to socioeconomic index for all in a zip code	CEPA

Seg_AA	segregation index between Caucasians and African-Americans – average deviation of each student's school ethnic diversity from the district-wide ethnic diversity; values of 0 indicate no segregation while values of 1 indicate complete segregation	CEPA
Seg_HI	same index as Seg_AA between Caucasians and Hispanics	CEPA

* NASBA is the National Association for Boards of Accountancy; DOE is the Department of Education; and CEPA is the Center for Education Policy Analysis at Stanford.

Table 3: Descriptive statistics

This table reports descriptive statistics for the sample of CPA candidates who took the exam between 2005 and 2016. The sample includes sittings by African-Americans, Hispanics, and Caucasians with complete data for variables of interest. Variables are defined in Table 2.

Variable	African-Americans (n=9,169)		Hispanics (n=10,076)		Caucasians (n=118,250)		Female (n=68,495)		Male (n=69,000)	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Pass_all_four_sections	0.201	0.401	0.316	0.465	0.470	0.499	0.410	0.492	0.472	0.499
Drop_after_first_attempt	0.197	0.398	0.153	0.360	0.080	0.271	0.102	0.303	0.084	0.278
Drop_after_first_section	0.047	0.211	0.039	0.193	0.020	0.141	0.025	0.157	0.022	0.145
Ethnicity_AA							0.079	0.270	0.055	0.227
Ethnicity_HI							0.081	0.274	0.065	0.247
Ethnicity_CA							0.840	0.367	0.880	0.325
Gender	0.589	0.492	0.554	0.497	0.486	0.500				
Age	32.609	8.301	29.837	6.829	28.511	6.943	29.303	7.562	28.463	6.609
Lapse	5.853	5.583	4.358	4.418	3.837	4.690	4.200	4.885	3.821	4.631
Num_attempts_total	5.090	5.024	5.070	4.375	5.159	3.807	5.280	4.183	5.018	3.685
First section attempted-AUD	0.277	0.447	0.257	0.437	0.264	0.441	0.268	0.443	0.261	0.439
First section attempted-BEC	0.242	0.428	0.227	0.419	0.215	0.411	0.219	0.413	0.217	0.412
First section attempted-FAR	0.300	0.458	0.328	0.470	0.336	0.472	0.325	0.468	0.342	0.474
First section attempted-REG	0.181	0.385	0.188	0.391	0.184	0.388	0.188	0.391	0.181	0.385
AACSB	0.651	0.477	0.748	0.434	0.756	0.430	0.732	0.443	0.764	0.425
Deg_IRS	0.957	0.202	0.977	0.151	0.987	0.113	0.980	0.141	0.989	0.105
Deg_High	0.204	0.403	0.147	0.354	0.149	0.356	0.147	0.354	0.158	0.365
School_High	0.888	0.315	0.932	0.252	0.896	0.305	0.897	0.304	0.900	0.300
Enrollment_High	0.608	0.488	0.723	0.447	0.653	0.476	0.644	0.479	0.666	0.472
Tuition_High	0.248	0.432	0.279	0.448	0.314	0.464	0.300	0.458	0.315	0.464
SatVerbal75	507.05	185.14	545.64	152.43	547.22	167.63	538.84	171.46	549.96	164.49
Pellperc	36.994	15.225	34.490	12.919	27.554	10.196	29.413	11.536	27.975	10.792
Ses_all	-0.146	0.939	0.033	0.884	0.331	0.895	0.246	0.891	0.309	0.922
Seg_AA	0.242	0.210	0.204	0.199	0.128	0.162	0.143	0.170	0.140	0.174
Seg_HI	0.189	0.146	0.161	0.138	0.106	0.120	0.117	0.125	0.115	0.127
Ethnicity_AA * Ses_AA	-8.796	94.035					-0.685	25.398	-0.489	23.327
Ethnicity_HI * Ses_HI			-4.135	44.743			-0.295	7.964	-0.311	15.220
Ethnicity_AA * Seg_AA	0.242	0.210					0.019	0.088	0.013	0.073
Ethnicity_HI * Seg_HI			0.161	0.138			0.013	0.058	0.011	0.054

Table 4: Spearman correlation coefficients

This table reports Spearman correlation coefficients for the sample of CPA candidates who took the exam between 2005-2016. Correlation coefficients in bold indicate a lack of significance at the 5% level. The sample includes sittings by Caucasians, African-Americans, and Hispanics with complete data for variables of interest. Variables are defined in Table 2.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Pass_all_four_sections																				
2. Drop_after_first_attempt	-0.28																			
3. Drop_after_first_section	-0.14	-0.05																		
4. Ethnicity_AA	-0.13	0.10	0.04																	
5. Ethnicity_HI	-0.07	0.06	0.03	-0.08																
6. Ethnicity_CA	0.15	-0.11	-0.05	-0.66	-0.70															
7. Gender	-0.06	0.03	0.01	0.05	0.03	-0.06														
8. Age	-0.22	0.08	0.07	0.14	0.06	-0.15	0.03													
9. Lapse	-0.14	-0.04	0.04	0.12	0.04	-0.12	0.05	0.74												
10. Num_attempts_total	0.36	-0.49	-0.16	-0.04	-0.03	0.05	0.02	0.09	0.26											
11. AACSB	0.13	-0.09	-0.04	-0.06	0.00	0.04	-0.04	-0.14	-0.08	0.04										
12. Deg_IRS	0.06	-0.06	-0.02	-0.06	-0.02	0.05	-0.04	-0.14	-0.05	0.04	0.22									
13. Deg_High	**0.00**	0.02	**0.00**	0.04	**0.00**	-0.02	-0.02	0.15	-0.06	-0.04	0.02	-0.08								
14. School_High	0.04	-0.02	**-0.00**	-0.01	0.03	-0.02	-0.01	0.01	-0.01	0.01	0.45	**0.00**	0.09							
15. Enrollment_High	0.10	-0.05	-0.03	-0.03	0.04	-0.01	-0.02	**-0.00**	-0.03	0.02	0.47	**-0.00**	0.04	0.41						
16. Tuition_High	-0.02	**0.00**	**0.00**	-0.03	-0.02	0.04	-0.02	-0.12	-0.05	0.01	-0.23	0.06	0.03	-0.25	-0.50					
17. SAT Verbal 75%	0.18	-0.11	-0.05	-0.09	-0.04	0.09	-0.05	-0.28	-0.15	0.05	0.30	0.20	-0.02	0.08	0.13	0.28				
18. Pell percentage (Pellperc)	-0.17	0.11	0.05	0.15	0.13	-0.21	0.06	0.35	0.21	-0.05	-0.27	-0.17	0.04	-0.04	-0.02	-0.31	-0.65			
19. Socio-economic index (Ses_all)	0.04	-0.03	-0.01	-0.12	-0.08	0.15	-0.04	-0.11	-0.09	0.01	**0.00**	0.01	-0.02	-0.02	-0.04	0.06	0.08	-0.19		
20. Segregation index for AA	**-0.00**	0.01	0.01	0.16	0.11	-0.20	0.02	0.11	0.12	**0.00**	0.07	-0.01	0.04	0.06	0.14	-0.07	0.09	0.07	-0.54	
21. Segregation index for HI	0.01	**0.00**	0.01	0.15	0.11	-0.19	0.02	0.12	0.11	0.01	0.08	-0.02	0.05	0.07	0.16	-0.06	0.12	0.05	-0.49	0.91

Table 5: Logistic regressions

This table reports the results of logistic regressions estimating the probability a candidate will pass all four sections (*Pass_all_four_sections = 1*), drop after first attempt (*Drop_after_first_attempt = 1*), or drop after first section (*Drop_after_first_section = 1*), as a function of candidate ethnicity, gender, and control variables:

$$
\begin{aligned}
&Prob\,(Pass\_all\_four\_sections, Drop\_after\_first\_attempt, Drop\_after\_first\_section)\\
&\quad = \beta_0 + \beta_1 Ethnicity\_AA + \beta_2 Ethnicity\_HI + \beta_3 Gender + \beta_4 Age + \beta_5 Lapse\\
&\quad + \beta_6 Num\_attempts\_total + \beta_7 First\_section\_attempted + \beta_8 AACSB\\
&\quad + \beta_9 Deg\_IRS + \beta_{10} Deg\_high + \beta_{11} School\_high + \beta_{12} Enrollmen\_high\\
&\quad + \beta_{13} Tuition\_high + \beta_{14} SatVerbal75 + \beta_{15} Pellperc + \beta_{16} Sesall\\
&\quad + \beta_{17} Hswhtblk + \beta_{18} Hswhthsp + \beta_{19} Ethnicity\_AA * Sesblk + \beta_{20} Ethnicity\_HI\\
&\quad * Seshsp + \beta_{21} Ethnicity\_AA * Hswhtblk + \beta_{22} Ethnicity\_HI * Hswhthsp\\
&\quad + \beta_{23} Year\_scheduled
\end{aligned}
$$

Variables are defined in Table 2. Model 2 excludes Num_attempts_total as it is a constant 1 for candidates who drop after the first attempt.

	Model 1 (n = 137,495)		**Model 2 (n = 137,495)**		**Model 3 (n = 137,495)**	
Variable	**Pass_all_ four_sect.**	**p-value**	**Drop_after_ first_attempt**	**p-value**	**Drop_after_ first_section**	**p-value**
Intercept	0.529	<.0001	-3.278	<.0001	-3.178	<.0001
Ethnicity_AA	0.459	<.0001	-0.433	<.0001	-0.086	0.048
Ethnicity_HI	0.192	<.0001	-0.273	<.0001	-0.158	0.000
Gender	-0.124	<.0001	0.060	<.0001	0.040	0.032
Age	-0.056	<.0001	0.048	<.0001	0.023	<.0001
Lapse	-0.019	<.0001	-0.065	<.0001	0.025	<.0001
Num_attempts_total	0.128	<.0001			-0.506	<.0001
First_section_attempted-AUD	-0.026	0.014	0.010	0.529	0.051	0.099
First_section_attempted-BEC	-0.230	<.0001	0.082	<.0001	0.100	0.002
First_section_attempted-FAR	0.263	<.0001	-0.023	0.131	-0.038	0.204
AACSB	-0.093	<.0001	0.111	<.0001	0.083	0.001
Deg_IRS	-0.077	0.014	0.034	0.272	-0.080	0.177
Deg_High	-0.173	<.0001	0.034	0.011	0.085	0.001
School_High	0.118	<.0001	-0.063	0.000	-0.157	<.0001
Enrollment_High	-0.151	<.0001	0.102	<.0001	0.097	<.0001
Tuition_High	0.121	<.0001	-0.058	<.0001	-0.073	0.004
SatVerbal75	0.001	<.0001	-0.000	<.0001	0.000	0.032
Pellperc	-0.018	<.0001	0.017	<.0001	0.011	<.0001
Ses_all	0.028	0.001	0.006	0.677	0.029	0.251
Seg_AA	-0.759	<.0001	0.629	<.0001	0.798	0.003
Seg_HI	1.261	<.0001	-0.814	<.0001	-0.549	0.119
Ethnicity_AA * Ses_AA	0.000	0.650	0.000	0.559	0.000	0.687
Ethnicity_HI * Ses_HI	-0.001	0.155	-0.000	0.715	0.000	0.845
Ethnicity_AA * Seg_AA	-0.441	0.003	-0.434	0.003	0.267	0.315
Ethnicity_HI * Seg_HI	-0.832	<.0001	0.234	0.294	-0.145	0.725
Wald Ratio	23843	<.0001	4882	<.0001	3130	<.0001

Table 6: Odds ratios and probabilities from logistic regressions

This table reports the odds ratios *(p/(1-p))* and probabilities *(p-value)* calculated based on the logit regressions shown as Equation (1) and reported in Table 5. Variables are defined in Table 2.

Variable	Model 1 (n = 137,495)		Model 2 (n = 137,495)		Model 3 (n = 137,495)	
	Pass_all_four_sect.	p-value	Drop_after_first_attempt	p-value	Drop_after_first_section	p-value
Ethnicity_AA	2.505	<.0001	0.421	<.0001	0.842	0.048
Ethnicity_HI	1.467	<.0001	0.579	<.0001	0.729	0.000
Gender	0.781	<.0001	1.127	<.0001	1.082	0.032
Age	0.946	<.0001	1.049	<.0001	1.023	<.0001
Lapse	0.981	<.0001	0.937	<.0001	1.025	<.0001
Num_attempts_total	1.136	<.0001			0.603	<.0001
First_section_attempted-AUD	0.981	0.299	1.082	0.007	1.178	0.004
First_section_attempted-BEC	0.800	<.0001	1.162	<.0001	1.237	0.000
First_section_attempted-FAR	1.309	<.0001	1.047	0.106	1.078	0.177
AACSB	0.830	<.0001	1.248	<.0001	1.181	0.001
Deg_IRS	0.857	0.014	1.070	0.272	0.852	0.177
Deg_High	0.708	<.0001	1.070	0.011	1.184	0.001
School_High	1.267	<.0001	0.882	0.000	0.730	<.0001
Enrollment_High	0.740	<.0001	1.227	<.0001	1.214	<.0001
Tuition_High	1.274	<.0001	0.891	<.0001	0.865	0.004
SatVerbal75	1.001	<.0001	1.000	<.0001	1.000	0.032
Pellperc	0.982	<.0001	1.017	<.0001	1.011	<.0001
Ses_all	1.029	0.001	1.006	0.677	1.030	0.251
Seg_AA	0.468	<.0001	1.876	<.0001	2.222	0.003
Seg_HI	3.530	<.0001	0.443	<.0001	0.578	0.119
Ethnicity_AA * Ses_AA	1.000	0.650	1.000	0.559	1.000	0.687
Ethnicity_HI * Ses_HI	0.999	0.155	1.000	0.715	1.000	0.845
Ethnicity_AA * Seg_AA	0.644	0.003	0.648	0.003	1.306	0.315
Ethnicity_HI * Seg_HI	0.435	<.0001	1.264	0.294	0.865	0.725

Table 7: Survival analysis

This table reports the results of Survival analysis estimating the probability a candidate will pass all four sections of CPA exam within the 18-month (548-day) window allowed. Both parametric and non-parametric test results are presented for robustness. Variables are defined in Table 2.

	Parametric (Cox proportional model)			Non-Parametric	
Variable	**Estimate**	**p-value**	**Hazard ratio - 1**	**Chi-Square**	**P-value**
Ethnicity_AA	0.511	<.0001	0.666	1968.1	<.0001
Ethnicity_HI	0.216	<.0001	0.241	661.6	<.0001
Gender	-0.119	<.0001	-0.112	552.9	<.0001
Age	-0.034	<.0001	-0.034	9568.2	<.0001
Lapse	-0.057	<.0001	-0.055	9131.3	<.0001
Num_attempts_total	-0.101	<.0001	-0.096	9831.8	<.0001
First_section_attempted-AUD	-0.028	0.025	-0.028	48.4	<.0001
First_section_attempted-BEC	-0.179	<.0001	-0.164	858.9	<.0001
First_section_attempted-FAR	0.168	<.0001	0.183	1272.7	<.0001
AACSB	-0.074	<.0001	-0.072	1593.7	<.0001
Deg_IRS	0.008	0.871	0.008	282.0	<.0001
Deg_High	-0.123	<.0001	-0.116	6.0	0.015
School_High	0.133	<.0001	0.142	79.8	<.0001
Enrollment_High	-0.157	<.0001	-0.146	1061.0	<.0001
Tuition_High	0.141	<.0001	0.152	46.1	<.0001
SatVerbal75	0.000	<.0001	0.000	1158.3	<.0001
Pellperc	-0.012	<.0001	-0.011	4266.7	<.0001
Ses_all	0.008	0.185	0.008	349.3	<.0001
Seg_AA	-0.316	<.0001	-0.271	240.3	<.0001
Seg_HI	0.320	<.0001	0.377	162.1	<.0001
Ethnicity_AA * Ses_AA	-0.000	0.055	0.000	9.7	0.002
Ethnicity_HI * Ses_HI	0.000	0.946	0.000	1.9	0.167
Ethnicity_AA * Seg_AA	-0.332	0.005	-0.282	1380.1	<.0001
Ethnicity_HI * Seg_HI	-0.527	0.000	-0.410	485.9	<.0001

Figure 1
Survival estimates – Ethnicity

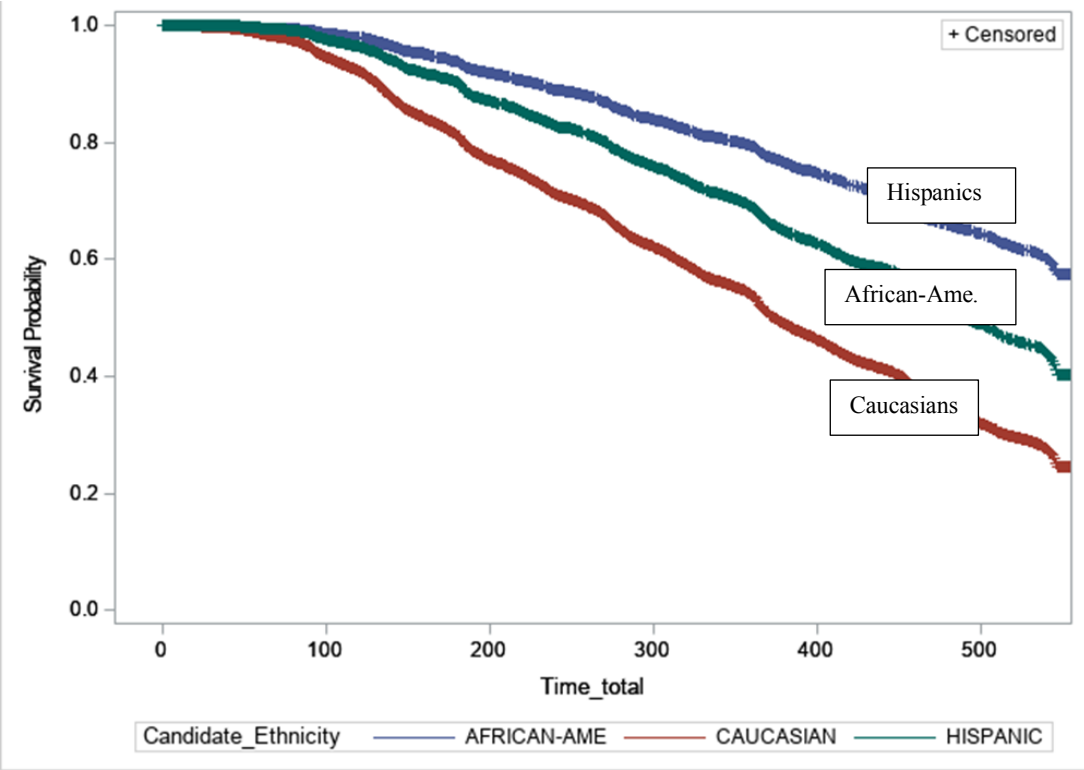

Figure 2
Survival estimates – Gender