




The Effects of Race/Ethnicity on Clinical Exam Outcomes:

Diminished (yet Persistent) Effects When Other Determinants Are Controlled

Prepared for
Association of Social Work Boards (ASWB)

Exam Report No. 3



Joy Kim, MSW, PhD & Michael Joo, MSW, PhD
Rutgers School of Social Work

Contents

Executive Summary	3
Background and Purpose.....	5
Conceptual Framework	6
Methods	7
Data and Sample	7
Variables and Measures	8
Examinees’ Demographic, Educational, and Employment Characteristics	8
Characteristics of MSW Programs	9
Characteristics of Educational Institutions	10
Community Characteristics.....	10
Data Analyses	11
Findings	12
Raw Scores	12
Pass Rate Disparities by Demographic Characteristics	15
Race/Ethnicity and Age Group.....	15
Gender and Primary Language	16
Pass Rate Disparities by Educational and Employment Backgrounds.....	17
Associate’s Degree and a BSW.....	17
Timing of the Exam.....	18
Years of Employment and Job Position.....	18
Pass Rate Disparities by Institutional Characteristics.....	20
MSW Program Size and Diversity.....	20
Type of Educational Institutions	21
Institution’s Location and Undergraduate Student’s Socioeconomic Status	22
Institution’s Undergraduate Admission Selectivity	23
Pass Rate Disparities by Community Characteristics and Region	24
Household Income and Racial Income Inequality	24
Region of Residence	26
Net Effect of Race/Ethnicity on the Odds Of Exam Failure	27
Discussion.....	29
Summary	29
Further Studies for Potential Interventions	30
Implications	31
References.....	33
Appendix Tables.....	35

Executive Summary

This report aims to investigate the determinants of the first-time ASWB Clinical exam passage and estimate the net effects of race/ethnicity on the exam outcome. This report — the third in the Exam Report Series — builds upon the second report, which reviewed the determinants of licensing exam disparities identified in other professions’ literature. The **conceptual framework** developed from the literature review was empirically tested with the ASWB Clinical exam data (1) to assess whether the determinants of ASWB exam passage are consistent with those in other licensed professions’ literature and (2) to estimate the net effects of race/ethnicity on the Clinical exam outcomes while holding the effects of other determinants constant. As the ASWB exam data provide only a limited number of variables on the examinees, the data were reinforced with zip code–level income data from the U.S. Census Bureau and institutional characteristics data from the U.S. Department of Education. The report begins with a closer look at the raw scores of Clinical examinees by race/ethnicity to better understand the disparate exam outcomes. **The analyses were confined to U.S. examinees who took the Clinical exam for the first time between 2018 and 2022.** Below are some of the key findings.

- ✓ First, the average score of all examinees was 110, and half of the examinees scored at least 111. When **raw scores** were compared to passing scores, Black examinees, on average, scored about four points below the passing scores. In comparison, examinees from other historically marginalized groups scored roughly three to six points above the passing scores.
- ✓ Second, exam outcomes were associated with race/ethnicity and most other demographic, institutional, and community characteristics examined in the analyses. Consistent with the existing evidence from other professions, Clinical exam pass rates varied significantly by **age group, gender, primary language, educational background, and employment experiences**. Those who began their postsecondary education with an associate’s degree and majored in social work as undergraduates had lower pass rates than their counterparts. Examinees who waited longer to take the exam after earning an MSW and had more years of employment also had lower pass rates than those who waited for a shorter period or worked for fewer years. Examinees who did not hold direct service positions had a lower pass rate.
- ✓ Third, as the literature suggested, the first-time Clinical exam pass rate was associated with the **characteristics of educational institutions that the examinees attended**. Those who attended smaller MSW programs, as well as institutions that were less selective in admission and mainly served students from lower socioeconomic backgrounds, had a lower pass rate than their counterparts.
- ✓ Last, statistical analyses suggested that if examinees from historically marginalized groups had the same demographic, educational, and employment characteristics and lived in similar institutional and community environments as white examinees, the Black–white disparity in the Clinical exam outcomes could be reduced by about 20%, and the Hispanic/Latino–white disparity by around 28%. **Black examinees’ exam outcomes were**

sensitive to institutional and community-level socioeconomic status and inequalities. On the other hand, the exam outcomes of Hispanic/Latino examinees were explained more by their demographic backgrounds.

Many professions have been challenged with racially disparate licensing exam outcomes. The prevalence suggests that the causes of disparities are deeply rooted in the fabric of our socioeconomic systems. In response to the disparities in exam outcomes, many take a reductionist approach by blaming the exams or advocating to remove competence assessment in the licensure system. While no licensing exam may be perfect as an assessment tool for professional competence in the complexity of real practice environments (Kane, 2005), group differences in exam outcomes do not necessarily indicate that the exams are biased. They instead reflect persistent inequalities and segregation in our schools, communities, and workplaces that disproportionately and adversely affect people from low-income and historically marginalized backgrounds (Hauser & Heubert, 1998).

The findings presented in this report should prompt many research questions and call for longer-term and more comprehensive empirical research that incorporates the crucial determinants of exam outcomes that this analysis could not incorporate due to data limitations. Assuming the causes of racially disparate exam outcomes are multifaceted, complex, and deeply rooted in our society, professional stakeholders must commit to collaborative research and strategic interventions to address the problem.

Background and Purpose

The *2022 ASWB Exam Pass Rate Analysis* revealed that racial/ethnic disparity in exam pass rates exists in the social work competence assessment. The analysis called for a further investigation to shed light on examinees' performance differences and, more importantly, the contributing factors. Having a deeper understanding of the level of disparity as well as its contributing factors is a critical first step toward developing effective interventions to reduce and eliminate the disparity. This Exam Report Series was conceived as a way to investigate the disparity and its determinants further.

The **first report** in the Exam Report Series presented the demographic, educational, and employment profiles of the most recent examinees of the Bachelors, Masters, and Clinical exams. The report highlighted that examinees from historically marginalized groups, particularly Black examinees, experienced delays in their social work education and training, which might have led them to take the licensing exams at older ages than others.

The **second report** provided a review of other licensed professions' literature. Other professions reported a similar level of disparate exam outcomes by race/ethnicity (e.g., Affrunti & Rossen, 2023; Rubright et al., 2019; Sharpless, 2021; Yeo et al., 2021). Empirical evidence from those professions supports that examinees' demographic and educational backgrounds, as well as the characteristics of their educational institutions and communities of residence, are the determinants of racially disparate exam outcomes (Espahbodi et al., 2023).

Applying the takeaways from other licensed professions' literature, **this third report shares findings from statistical analyses of the ASWB's Clinical exam data to answer the following research questions.**

- 1) First, how do the ASWB examinees' performance — measured in raw scores — differ by race/ethnicity? Although a licensing exam is designed to discern a professional candidate's competence through a pass/fail outcome, raw score analyses provide additional information as to how examinees' performances are different by race/ethnicity.
- 2) Second, what demographic, educational, and employment characteristics of examinees, as well as the characteristics of their educational institutions and communities, are significantly associated with their ASWB exam outcomes? What are the positive or negative predictors of ASWB exam passage? Are the factors of ASWB exam passage consistent with the determinants identified in other professions' literature reviewed in Exam Report No. 2?
- 3) Third, are the negative predictors of ASWB exam passage, such as delayed exam taking, more prevalent among examinees from historically marginalized backgrounds?
- 4) Last, what is the estimated net effect of race/ethnicity on ASWB exam passage, holding the effects of other factors constant? How does the effect of race/ethnicity decline when other determinants of exam outcomes are considered in the analyses?

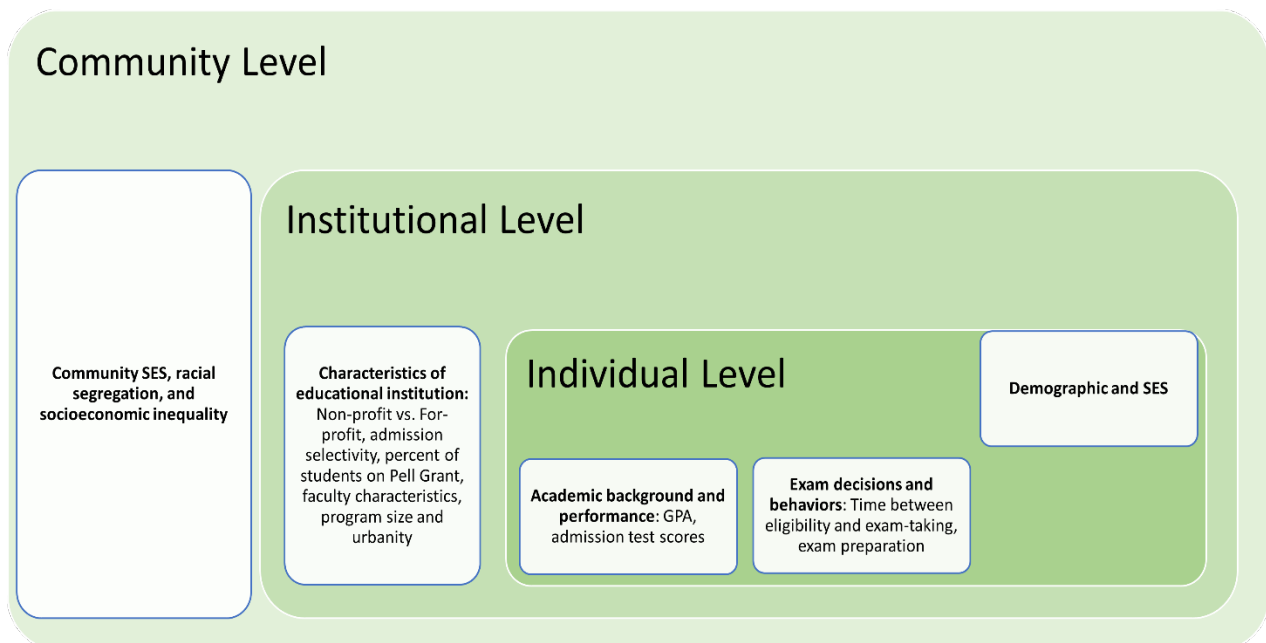
The following section discusses a conceptual framework and methodologies used to answer these four questions through statistical analyses.

Conceptual Framework

Figure 1 presents the determinants of licensing exam outcomes at individual, institutional, and community levels, as discussed in Exam Report No. 2. It serves as a conceptual framework that guides the empirical analyses of pass rate disparities. It depicts how examinees' exam outcomes are influenced not only by their demographic and socioeconomic backgrounds but also by the characteristics of their educational institutions and the broader socioeconomic opportunities and disadvantages in their communities.

Figure 1 also illustrates the data necessary to investigate licensing exam outcomes. To empirically test the framework depicted in the figure, one would need data that capture the examinees' socioeconomic status, detailed academic background and performance, and exam-related behaviors (such as how long after obtaining MSWs they took the exam, how they met the qualified supervision requirements, and how they prepared for the exam). Also necessary are data on the detailed characteristics of MSW programs that the examinees attended, such as admission selectivity, curriculum, faculty characteristics, and program size. Additionally, one would need data that measure the socioeconomic characteristics of the communities where the examinees lived, including household incomes and indicators of racial segregation and inequality.

Figure 1. Factors That Affect Licensing Exam Outcomes



As the next section discusses, only some of the necessary data were available in the ASWB exam data file for statistical analyses. The data limitation was a barrier to comprehensive statistical testing of the conceptual framework for social work.

Methods

Data and Sample

The primary data for this analysis came from the **Clinical exam data file** provided by ASWB. The sample was restricted to 88,678 first-time examinees in the United States between 2018 and 2022. Repeat examinees were excluded from the analysis as the examinees' performance is best described with the first-time scores and/or pass/fail outcomes (National Council on Teacher Quality, 2019). The year 2018 was chosen because it was the first year when the most recent exam blueprint was used. Because a single-year data file did not provide a sufficient number of Asian examinees, five years of data between 2018 and 2022 were pulled together to increase the sample size.

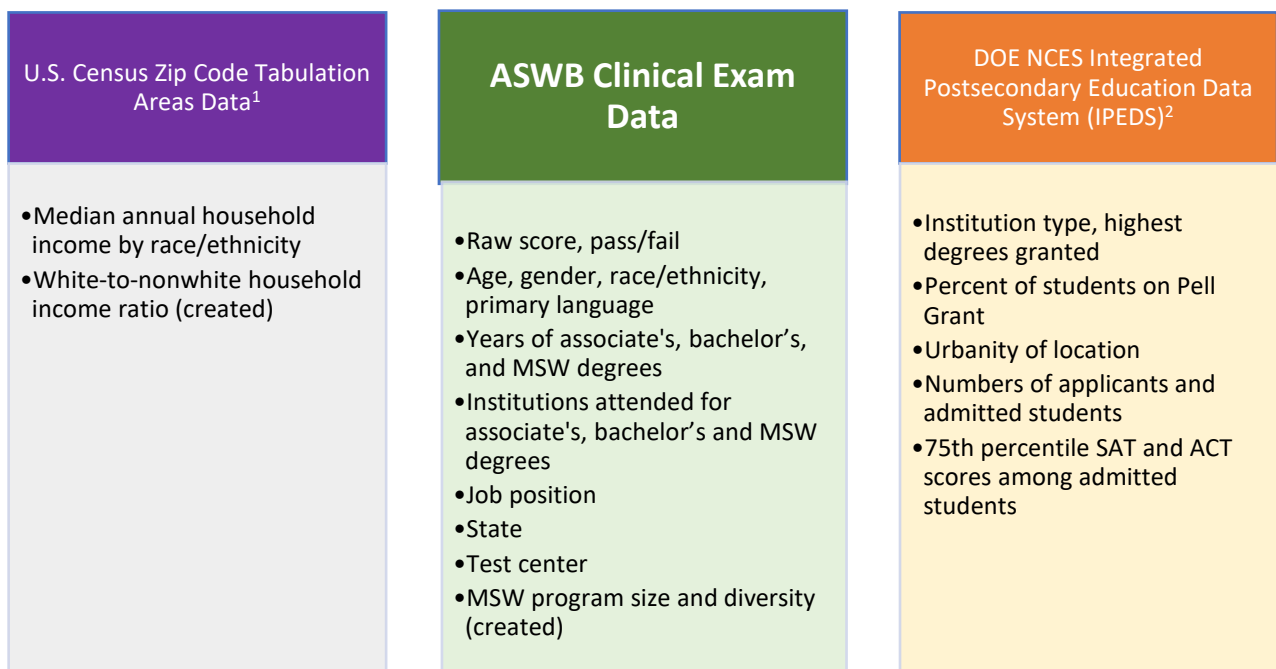
Social work licensing examinees' demographic data were collected when examinees registered for the exams. The available demographic characteristics of Clinical examinees were age, gender, race/ethnicity, English use as the primary language, and the state of residence. The available educational characteristics of the examinees were the years of associate's, bachelor's, and master's degrees and the identifiers and state locations of their educational institutions. Whether or not examinees had majored in social work as undergraduates was also available. Some of the examinees' employment characteristics, including years of employment and job positions, were available.

While these data can serve as important determinants of their exam passage, testing the conceptual framework depicted in Figure 1 required much more individual data as well as data for examinees' educational institutions and communities of residence. The necessary data include examinees' academic backgrounds and performance (e.g., GPAs in their academic programs and admission test scores), the characteristics of MSW programs that they attended (e.g., faculty characteristics, admission selectivity, curriculum, etc.), and the socioeconomic characteristics of the communities where the examinees grew up and lived (e.g., household incomes, racial/ethnic income inequality). **Unfortunately, the ASWB exam data file did not include any of them.**

To remedy these data limitations to the extent possible, the ASWB Clinical exam data were merged with data from the U.S. Census ZCTA tabulation areas and the U.S. Department of Education's National Center for Education Statistics (NCES) Integrated Postsecondary Education Data System (IPEDS) as shown in Figure 2. The Census data merge was to get **zip code-level household income data**, and the IPEDS data merge was to obtain public data about **institutions' characteristics**. The ASWB exam data contained two important variables to allow the data files

to merge: (1) the IDs of the test centers where exams took place and (2) the IDs of the educational institutions that the examinees had attended for graduate degrees. We received the test centers' zip codes from ASWB. We then obtained zip code–based median annual household incomes (between 2018 and 2022) by race/ethnicity from the U.S. Census. We merged the census data file with the exam data file, using the zip codes as the common denominator. Additionally, the IDs of the colleges or universities that the examinees had attended for their graduate degrees were identified in the institutional IDs of the 2018 NCES IPEDS data. Using the institutional IDs as the common denominator, the exam data file and the IPEDS data file were merged. Figure 2 summarizes the source files of all the variables included in the statistical analyses that will be discussed below.

Figure 2. Three Data Sources for This Analysis



1. <https://data.census.gov/table>

2. <https://nces.ed.gov/ipeds/datacenter/InstitutionByName.aspx?goToReportId=5&sid=3ffcc66-2217-442c-b475-bfa5c5dcaa52&rtid=5>

Variables and Measures

Examinees' Demographic, Educational, and Employment Characteristics

Clinical examinees' basic demographic characteristics were available in the ASWB exam data. The **race/ethnicity** variable recorded if examinees were Asian, Black, Hispanic/Latino, multiracial, **Native American/Indigenous Peoples**, or white. Note that some examinees (less than 2%) did not disclose their race/ethnicity and were categorized as 'unknown' in the analysis. **In addition, the number of Native American/Indigenous Peoples examinees was too small for**

analyses; data for that group was not used in this report. Using the age variable, the following four **age categories** were created: (1) age 29 or below, (2) between 30 and 34, (3) between 35 and 39, and (4) 40 or older. **Gender** was categorized as male or female. (Because less than 0.1% reported being a gender not listed, they could not be included in the analyses.) The examinees' **English use** was described as primary use or secondary use.

The exam data also documented whether the examinees held an **associate's degree** or not and had **majored in social work** for their undergraduate degrees. The exam data also contained the years when the examinees obtained their educational degrees. Using the years when they earned their master's degrees, we created a variable that measured the years between the examinees' master's degrees and their first Clinical exam attempt. The variable was labeled **years since MSW** and was categorized into the following five groups: (1) less than or equal to one year, (2) between one and two years, (3) between three and four years, (4) between five and six years, and (5) more than six years. Although Clinical exam candidates are required to complete postgraduate supervised clinical training hours that typically take a couple of years, some examinees with less than one-year post-MSW were shown in the exam data (In some states, early approval of the exam might have been available.) The exam data file also had a variable that described examinees' **years of employment**, which was categorized into the following five groups: (1) less than or equal to one year, (2) between two and three years, (3) between four and five years, (4) between six and nine years, and (5) ten years or longer. The available **job position** variable measured if the examinees worked in direct service positions, administrative positions, other positions, or did not work (labeled as not applicable).

As discussed earlier, the exam data did not have more detailed information about the examinees' academic and socioeconomic backgrounds, such as examinees' GPAs in their academic programs, admission test scores, parents' education, and household incomes (e.g., Nettles, 2011, Rubright et al., 2019; Wightman, 1998). Lacking such important information limited the scope of our statistical analysis.

Characteristics of MSW Programs

The characteristics of MSW programs that the examinees had attended (e.g., faculty characteristics, admission selectivity, curriculum) are important determinants of their exam outcomes (e.g., Bline et al., 2016; Chaparro, 2020). However, the only relevant information available in the exam data file was the IDs of the institutions that the examinees had attended for their graduate degrees. To address this data limitation to the extent possible, we created two variables that could work as crude proxies of (1) the **size of MSW programs** and (2) the racial/ethnic **diversity of student bodies**. Using the **five-year total number of Clinical examinees** by educational institution, we created a variable that could indicate the size of MSW programs in the following categories based on the percentile distribution of the variable: (1) fewer than 506 examinees (smallest program), (2) between 507 and 1,097 examinees, (3) between 1,098 and 1,934 examinees, and (4) 1,935 examinees or more (largest program). To estimate the racial/ethnic diversity of MSW programs, we created a ratio variable, "examinees of color-to-white examinees." The ratio variable was created using the number of

examinees from historically marginalized groups and the number of white examinees by the educational institution. It had four categories as follows: (1) less than 24% examinees from historically marginalized groups (least diverse), (2) between 24% and 33%, (3) between 33% and 48%, and (4) more than 48% (most diverse).

Characteristics of Educational Institutions

The NCES IPEDS data provided the following eight variables to describe the characteristics of educational institutions that the examinees had attended for their MSWs. Many previous studies identified them as contributing factors to licensing exam outcomes (e.g., Angelo et al., 2021; Espahbodi et al., 2023; Falcone, 2012; Trinkle et al., 2016; Mittlestaedt & Morris, 2017). First, the **type of educational institutions** measured if an institution was private for-profit, private not-for-profit, or public. Second, the **highest degrees granted** by an institution included the following four categories: (1) a master's, bachelor's, or associate's degree, (2) doctoral degrees including research, (3) doctoral degrees including professional practice, and (4) doctoral degrees including research and professional practice. Third, the **urbanity of an institution** described if an institution was located in a rural area, suburban area, small city, midsize city, or large city. Fourth, the **percentage of undergraduate students on the Pell Grant** was categorized into (1) less than 20%, (2) between 20% and 31%, (3) between 32% and 49%, and (4) greater than 49%. Fifth, the **percentage of undergraduate applicants admitted** to the institutions was categorized into (1) less than 46%, (2) between 46% and 63%, (3) between 64% and 75%, and (4) greater than 75%. Sixth, the 75th percentile **SAT reading scores** of admitted first-year undergraduate students were grouped into (1) less than 600, (2) between 600 and 639, (3) between 640 and 690, and (4) above 690. Seventh, the 75th percentile **SAT math scores** were grouped into (1) less than 590, (2) between 590 and 639, (3) between 640 and 710, and (4) greater than 710. Last, the 75th percentile **ACT composite scores** of admitted undergraduate students were recorded as (1) less than 25, (2) between 25 and 27, (3) between 28 and 31, and (4) above 31. Note that some examinees did not have these scores as their institutions did not report admitted undergraduate students' SAT and ACT scores. Furthermore, it is important to note that these broad institutional characteristics measured with undergraduate student bodies did not measure the school characteristics and admission standards of MSW programs attended by the Clinical examinees.

Community Characteristics

The socioeconomic characteristics of the communities where the examinees grew up and lived are important determinants of exam outcomes (Espahbodi et al., 2023), but the exam data file did not have any relevant variables. Again, to address the data limitation, the zip code-level median household incomes of test centers where the examinees took the exams were used as **crude proxies of their income backgrounds**. However, as there were only 270 test center locations for more than 88,000 examinees, it is important to acknowledge that using test centers as examinees' geographic locations is likely to reduce true variations in the examinees' socioeconomic statuses severely. Despite the concern, we proceeded with creating the variable due to the lack of any alternative. The zip code-level five-year average **median annual**

household incomes were recorded as (1) less than \$57,665, (2) between \$57,665 and \$72,133, (3) between \$72,134 and \$92,881, and (4) greater than \$92,881 based on a percentile distribution of the data. In addition, using the zip code–level median annual household incomes by race/ethnicity, we created an “examinees of color-to-white examinees” income ratio to **measure racial/ethnic income inequality** by zip code. The income ratio was recorded as (1) less than 0.61 (most unequal), (2) between 0.61 and 0.76, (3) between 0.77 and 0.94, and (4) greater than 0.95 (about equal) based on a percentile distribution of the data.

The U.S. Census zip code data did not have income data for a multiracial group. In addition, income data for the examinees whose race/ethnicity was unknown could not be identified. Besides, not all zip codes had income data for Black, Hispanic/Latino, and Asian groups, as there were areas with insufficient residents from those groups. Therefore, some examinees’ income and inequality variables were missing and could not be included in the statistical analyses presented below.

Furthermore, using the states that authorized their exam registration, we grouped the examinees by the nine **regions of residence** as follows: South Atlantic, East North Central, East South Central, West South Central, Middle Atlantic, Mountain, Pacific, New England, and West North Central regions.

Data Analyses

First, we present detailed analyses of exam scores and pass rates (which are presented in Table 1 and Chart 1). We also present descriptive statistics on the examinees to understand their demographic, educational, and employment characteristics as well as the characteristics of their educational institutions and communities. These descriptive statistics are presented in **Appendix Table A-1**.

Then, we proceeded to examine the **relationship** between demographic, institutional, and community determinants and exam passage. Tukey's multiple comparison tests were conducted in SAS 9.4 to determine which means amongst a set of means differ from the rest. These results are presented in Charts 2 through 22. In describing the relationships, we used the word *significant* to mean statistical significance at least at $p < 0.05$, meaning that there is less than a 5% chance of obtaining the result by chance and more than a 95% chance that the result reflects a true relationship or difference in the population under study. Where possible, we connected the findings of our analyses to the literature of other professions to assess how our findings were congruent with the existing knowledge about licensing exam outcomes.

Finally, we ran a logistic regression analysis to examine **the net effect of race/ethnicity on exam failure** while controlling for other determinants/predictors of the exam outcome, including examinees’ characteristics and the characteristics of their institutions and communities. A logistic regression is a statistical method designed to examine associations between predictor variables (e.g., the characteristics of examinees and their institutions and communities) and a dichotomously measured outcome variable of interest (e.g., exam

pass/fail). We chose to predict exam failure rather than passage because we intended to examine how the negative effects of being a member of a historically marginalized group could be **reduced** when other predictors were considered in the analyses. Predicting failure instead of passage made the narration of the findings easier. As stated earlier, because the multiracial and unknown race/ethnicity groups were missing household income and racial inequality variables, those two groups were excluded from the regression analyses. The findings of our logistic regression analyses are summarized in Table 2, and the detailed findings are presented in **Appendix Table A-2**.

In building the logistic regression models, only the variables significantly related to exam failure were included. Also, as many predictor variables were highly correlated with one another (e.g., correlations between the 75th percentile SAT reading and math scores, correlations between zip code–level incomes and racial income ratios), we used caution in building a model to avoid high correlations among the variables by including only the necessary predictors that generated the best model fit statistics. Most importantly, to observe how the effects of the race/ethnicity variable changed with the inclusion of other predictor variables, we first ran a base model with only the race/ethnicity variable and then added other demographic, educational, and employment characteristic variables to the model. In the final model, we added the institutional and zip code–level community variables to examine how the additions changed the effects of race/ethnicity on the likelihood of exam failure. Our focus was to examine if and to what extent the effect of race/ethnicity was reduced due to the added other determinants of exam outcomes.

When interpreting the findings of the regression analyses, it would be important to note that the relationships explored in the analyses were **not causal, but only correlational**.

Findings

A detailed raw score analysis of the Clinical exam data revealed that on average, Black examinees scored around four points below the passing scores, but examinees from other historically marginalized groups scored roughly three to six points above the passing scores.

Raw Scores

Before examining exam pass rates, raw scores and their distributions were examined to gain a more thorough understanding of the exam outcomes. Table 1 shows the average raw scores, scores at percentile ranks, maximum score, and standard deviation of the scores for all examinees and by race/ethnicity. The average score for all examinees was 110, with a minimum

score of 25 and a maximum score of 143. The median (the 50th percentile) score of 111 suggests that one-half of all examinees scored at least 111. The table also suggests that half of all examinees scored between 102 (the 25th percentile score) and 119 (the 75th percentile score). The 5th and 10th percentile scores of all examinees were 85 and 92, respectively.

The table shows significant racial/ethnic disparities in raw scores, particularly between Black and white examinees. **Black examinees' average raw score was 99, about 14 points — roughly equivalent to one standard deviation — below white examinees' average score, which was 113.** Hispanic/Latino examinees' average score was 106, about seven points below white examinees' score. The scores of Asian, multiracial, and unknown racial groups were similar to one another at around 108–109.

The raw scores at the percentile rank reveal more important details about how scores were distributed differently by race/ethnicity. White examinees' scores surpassed the scores of examinees from other racial/ethnic groups at each percentile rank reported in Table 1. Yet, the score differences were rather modest, particularly at the 90th percentile ranks, where examinees from all racial/ethnic groups scored above 120, except for Black examinees. Interestingly, the Black–white score difference was the greatest at every percentile rank. For example, for the lowest performing (1st percentile) group, the Black–white score difference was 60 versus 81, and for the highest performing (90th percentile) group, the difference was 116 versus 127.

Looking at the mean difference from a passing score by race/ethnicity, Black examinees were the only group with a difference of -3.79. This suggests that **Black examinees, on average, scored around four points below the passing scores, but examinees from other marginalized groups scored roughly three to six points above the passing scores.**

Table 1. Clinical Exam Raw Score by Race/Ethnicity (2018-2022)

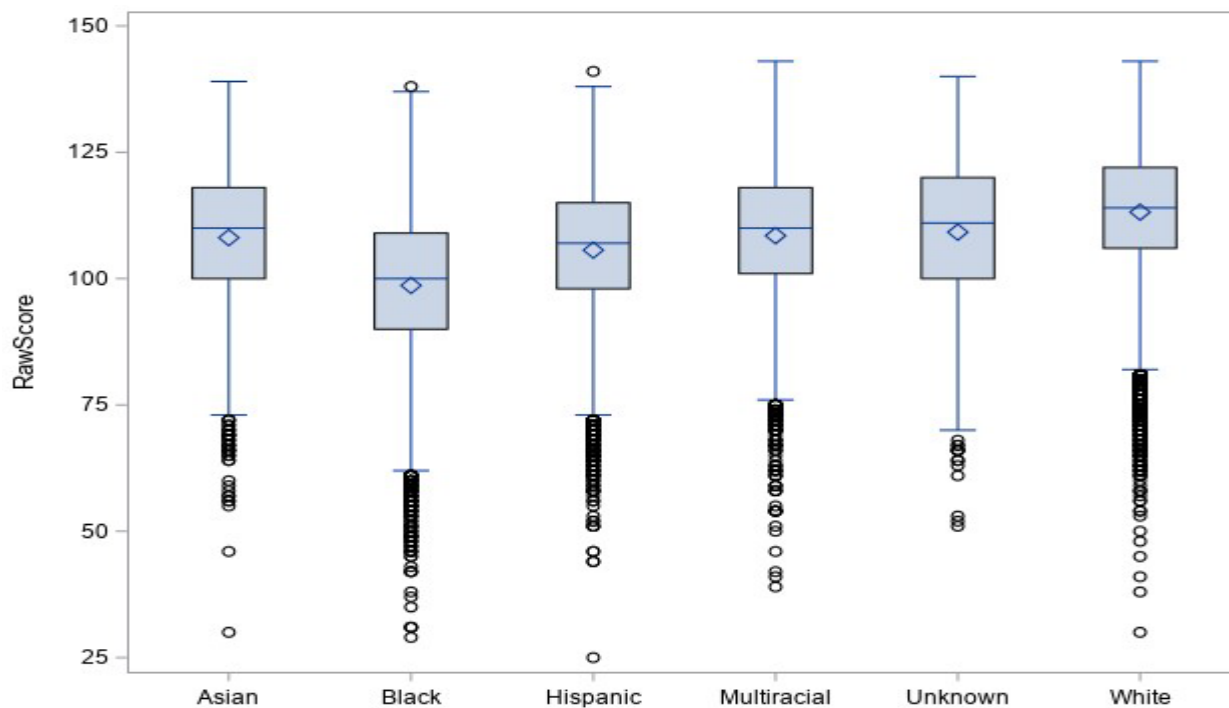
	All	Asian	Black	Hispanic/Latino	Multiracial	Unknown	White
<i>N</i>	88,678	3,146	12,530	10,572	4,714	1,373	56,343
Mean score	110	108	99	106	108	109	113
Score at percentile							
1 st	72	69	60	69	70	68	81
5 th	85	84	74	83	83	84	92
10 th	92	90	80	89	91	91	97
25 th	102	100	90	98	101	100	106
50 th	111	110	100	107	110	111	114
75 th	119	118	109	115	118	120	122
90 th	126	124	116	121	125	126	127
Maximum score	143	139	138	141	143	140	143
Standard deviation	13.56	13.38	14.21	13.02	13.81	14.13	11.80

The mean difference from a passing score	7.31	5.67	-3.79	3.22	6.08	6.80	10.76
--	------	------	-------	------	------	------	-------

Chart 1 provides additional visualizations of the raw score distributions by race/ethnicity. Using the 25th and 75th percentile scores in Table 1, Chart 1 illustrates where 50% of the scores were found in the distributions and how the distribution differed by race/ethnicity. Please note that the horizontal lines that split the boxes in two indicate the median scores and the diamonds indicate the mean scores. Overall, Chart 1 confirms that racial/ethnic disparities exist in median and mean scores and that Black examinees' median and mean scores were the lowest of all groups.

In Chart 1, the upper line stretching outside each box indicates the 75th percentile score to the maximum raw score for each race/ethnicity. Conversely, the lower line outside each box marks the 25th percentile score to the minimum raw score for each race/ethnicity. According to the plot, **white examinees' scores were relatively tightly distributed above 80. The scores of examinees from other historically marginalized groups — particularly Black examinees' scores — were more dispersed to include scores even below 70.** In addition, the circles outside the minimum scores in Chart 1 suggest that many extremely low scores existed as potential outliers, especially among Black examinees.

Chart 1. Box Plot of Clinical Exam Raw Scores by Race/Ethnicity, 2018–2022



Pass Rate Disparities by Demographic Characteristics

Examinees from historically marginalized groups, older groups, men, and those who use English as a second language had lower pass rates than their counterparts. These demographic disparities were consistent with findings from other professions' licensing exams.

Race/Ethnicity and Age Group

Analyses suggested that Clinical exam pass rates were significantly lower for Asian, Black, Hispanic/Latino, and multiracial groups than for white examinees. Chart 2 depicts the pass rates by race/ethnicity. The rates for Asian (0.71), Black (0.44), Hispanic/Latino (0.64), multiracial (0.72), and examinees whose race/ethnicity was not reported (0.71) were lower than the rate for white examinees (0.83). This significant racial disparity in exam pass rate is consistent with the findings from other regulated professions (Nettles et al., 2011; Rubright et al., 2019; Wightman, 1998; Yeo et al., 2020).

Analyses also found that the pass rates differed significantly by age group. As shown in Chart 3, older examinees had lower pass rates than younger examinees. While 80% of examinees in their 20s passed the exam, 66% of examinees in their 40s passed the exam. Age disparities in licensing exam pass rates are well documented in the literature (Blaine et al., 2016; Mitttestaedt & Morris, 2017; Nayer & Grover Takahashi, 2017; Nguyen et al., 2021; Trinkle et al., 2016), and this finding is aligned with previous studies in other professions.

Chart 2. Pass Rate by Race/Ethnicity, 2018-2022

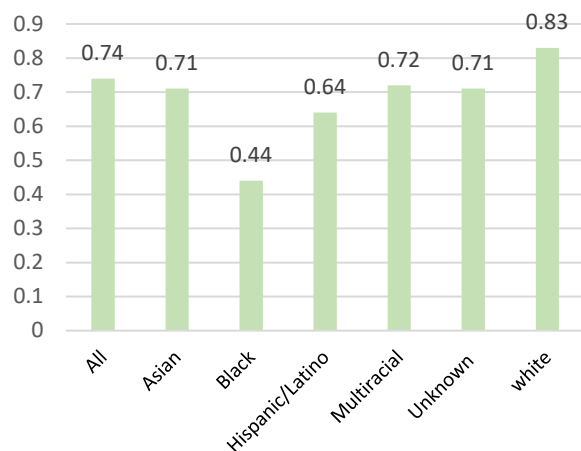
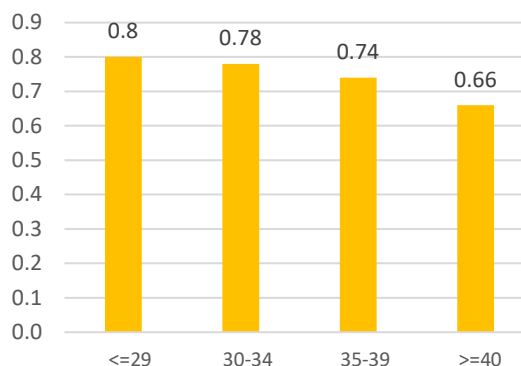


Chart 3. Pass Rate By Age Group, 2018-2022

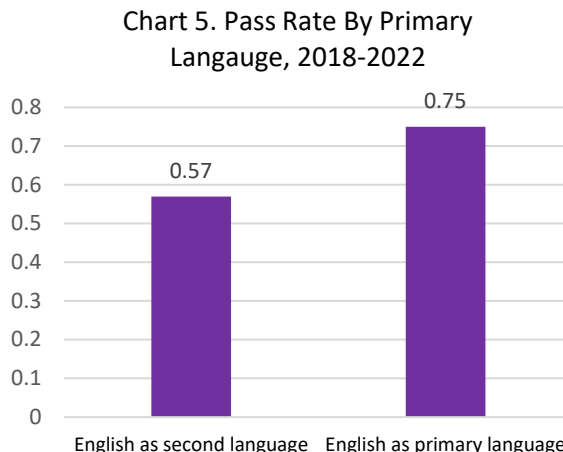
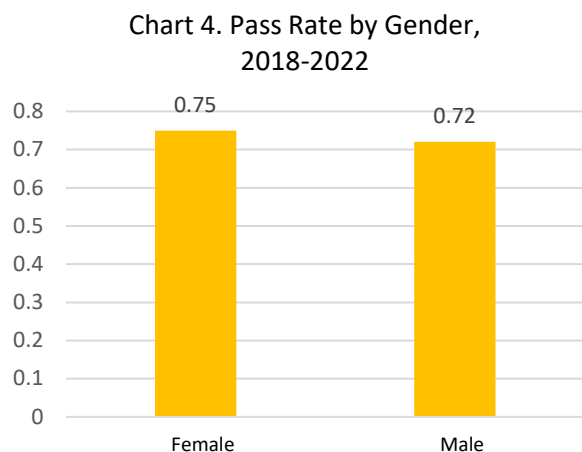


The bivariate analyses showed that older examinees, particularly those in their 40s, were more likely to fail the exam. **Are examinees from historically marginalized groups significantly older than their white counterparts?** Findings presented in **Appendix Table A-1** show the distribution of age groups by race/ethnicity. The table suggests that a higher percentage of Black examinees (37.26%) and examinees in the unknown race/ethnicity group (38.97) were in their 40s than white examinees (28.87%). The findings suggest that historically marginalized groups are significantly older than white examinees among Clinical examinees.

Gender and Primary Language

As Chart 4 shows, the exam pass rates significantly varied by gender. Women had a higher pass rate than men (0.75 versus 0.72). Although previous studies have reported gender disparity in licensing exam pass rates (Trinkle et al., 2016; Yeo et al., 2020), the ASWB exam appears unique in that women, not men, had a higher pass rate.

Chart 5 shows pass rate disparity by English use. Examinees who used English as a second language had a significantly lower pass rate than those whose primary language was English (0.57 versus 0.75). The use of English as a secondary language typically signifies the examinees' immigration status from a non-English speaking country. The disparity is similar to findings from other licensed professions that investigated pass rate differences by domestic and international examinees (Eich & O'Neill, 2017).



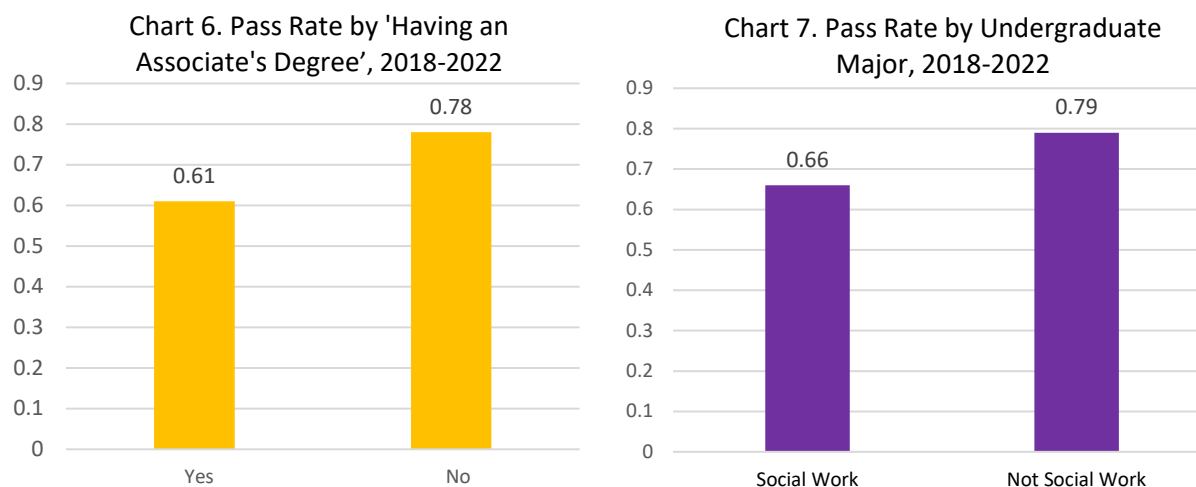
Are these negative predictors of exam outcomes more concentrated in historically marginalized groups of examinees? Interestingly, according to **Appendix Table A-1**, a higher percentage of the unknown race/ethnicity group than white examinees was male (18.94% versus 12.70%). In addition, nearly 33% of Asian and Hispanic/Latino examinees used English as a second language, compared to 3.63% of Black examinees or 1.36% of white examinees. Among multiracial and unknown race/ethnicity examinees, 9.44% and 6.26% used English as a second language. So, the negative predictors of exam outcome were more prevalent among members of historically marginalized groups.

Pass Rate Disparities by Educational and Employment Backgrounds

Clinical examinees who held an associate's degree and a BSW had lower pass rates. Those who waited longer to take the first Clinical exam after earning an MSW and had more years of employment had lower pass rates. Examinees who held non-direct service jobs also had lower pass rates.

Associate's Degree and a BSW

As Chart 6 shows, beginning postsecondary education with an associate's degree was negatively related to Clinical exam pass rates. Examinees with an associate's degree had a pass rate of 0.61, significantly lower than the rate of 0.78 for those who began their postsecondary education with a four-year degree. Undergraduate major was also related to exam pass rates. As shown in Chart 7, examinees who majored in social work for their undergraduate degrees had a pass rate of 0.66, significantly lower than 0.79 for those who had other undergraduate majors. This finding on the relationship between BSWs and Clinical exam pass rate seems counterintuitive. Yet, Nettles et al. (2011), who examined teacher licensing exam outcomes in Praxis I (Mathematics), reported the similar finding that education majors had a lower pass rate on the Praxis exam than non-education majors (Nettles et al., 2011).

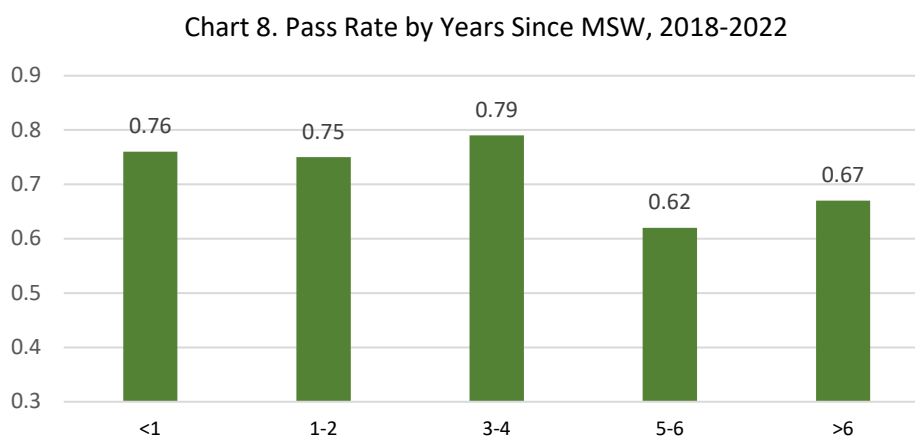


Beginning postsecondary education with an associate's degree and majoring in social work as an undergraduate were negatively related to Clinical exam passage. **Were these negative predictors of exam outcomes more prevalent among examinees from historically marginalized groups?** Appendix Table A-1 presents that a significantly higher percentage of Hispanic/Latino (30.34%), multiracial (25.58%), and Black examinees (24.03%) began their postsecondary education with an associate's degree than white examinees (18.61%). At the same time, more than 40% of Black examinees were social work majors, compared to 33.27% of

white examinees. That is, a higher share of historically marginalized race/ethnic groups carried educational characteristics that were negatively associated with exam outcomes.

Timing of the Exam

Chart 8 shows that the exam pass rate was significantly related to the number of years between when examinees earned their MSWs and when they took their first exams. Examinees who took the Clinical exam for the first time more than five years after they obtained their MSWs had a significantly lower pass rate than those who did it within three to four years after their MSWs (0.62 versus 0.79). Previous studies similarly reported that delayed exam-taking is negatively related to exam passage (Eich & O’Neill, 2017; Espahbodi et al., 2023; Nettles et al, 2011).



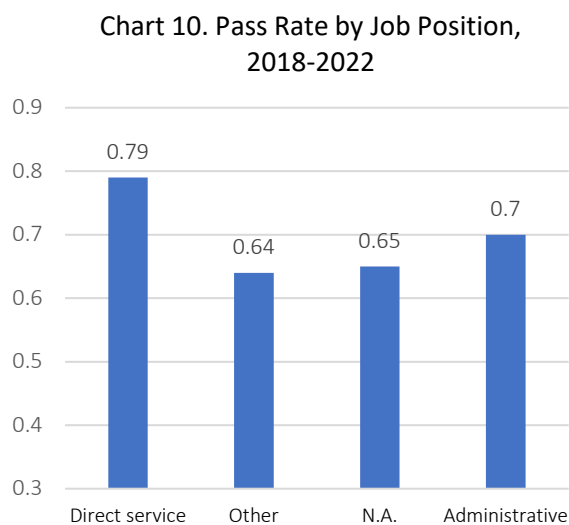
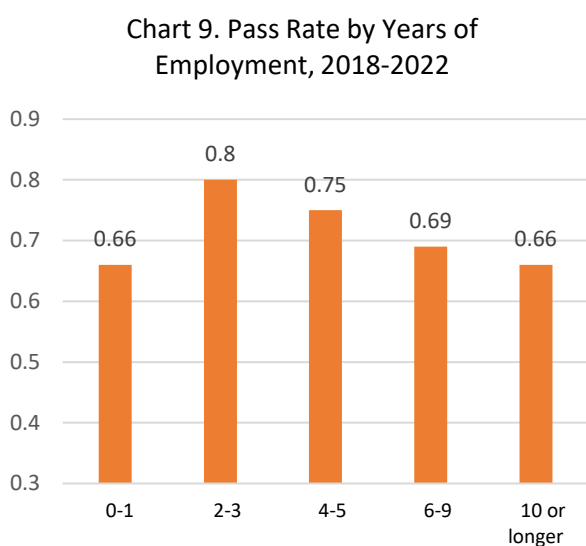
As with other negative predictors of exam outcomes, delayed exam-taking was more prevalent among examinees from historically marginalized groups. **Appendix Table A-1** suggested that a much higher percentage of Black examinees (nearly 53%), compared to about 37% of white examinees, took their first Clinical exam at least five years after earning their MSWs. Examinees from other historically marginalized groups also showed delays in exam-taking relative to their white counterparts. The exam delays may be related to challenges in meeting the required clinical supervision hours and warrant further studies.

Years of Employment and Job Position

The exam pass rate was significantly associated with years of employment. As Chart 9 suggests, the pass rate was the highest at 0.80 among examinees who had about two to three years of employment, and the rate was the lowest at 0.66 among those with at least 10 years of employment or less than a year of employment.

Job position was also significantly related to the exam pass rates. Examinees with direct service positions had the highest pass rate of 0.79, compared to those holding administrative positions (0.70) or “other” positions (0.64). It is possible that having a direct service position,

relative to other positions, is more likely to facilitate the development of clinical social work competence. According to the NASW/ASWB's national guidelines for clinical social work supervision (NASW & ASWB, 2013), candidates for clinical social worker licensure should complete the required supervised training "in an appropriate setting" to be eligible for the Clinical exam. In addition, the ASWB policy manual on the examinations (2022) states that the Clinical exam is developed for candidates with two years of experience in "clinical settings." The exam data suggest that some examinees held positions that might not have been conducive to developing clinical social work competence compared to those who held a direct service position.



Were these negative predictors of exam outcomes more prevalent among examinees from historically marginalized groups? Again, **Appendix Table A-1** suggests that relative to about 21% of white examinees who had more than six years of employment at their first exam attempt, nearly 32% of Black examinees had more than six years of employment. In addition, the lowest share of Black examinees (54%), compared to other racial/ethnic groups, had direct service positions. On the other hand, more than 65% of white examinees reported having a direct service position. The difference in job positions by race/ethnicity may indicate race/ethnicity-specific experiences in the social work labor market, which also warrants further studies.

Pass Rate Disparities by Institutional Characteristics

The ASWB Clinical exam pass rate was associated with the characteristics of MSW programs that the examinees attended. Those who graduated from large programs and programs with fewer students from historically marginalized groups had a higher pass rate than their counterparts.

The characteristics of educational institutions were related to exam outcomes. Clinical examinees who attended institutions that were less selective in admission and where more than half of students were Pell Grant recipients had a lower pass rate than their counterparts.

MSW Program Size and Diversity

As Chart 11 shows, examinees who graduated from the smallest programs (with less than 506 examinees) had the lowest pass rate of 0.69, and those who graduated from the largest (with at least 1,935 examinees) had the highest pass rate of 0.78. As indicated by the literature, the programs with a large number of Clinical examinees might have had more resources to aid examinees' preparation for the exam (Falcone, 2012).

Examinees who attended MSW programs in which up to a third (33%) of the students were from historically marginalized groups had a pass rate of around 0.80 but those who had attended programs in which nearly half of the students were from historically marginalized groups had the lowest pass rate of 0.65.

Chart 11. Pass Rate By Program Size, 2018-2022

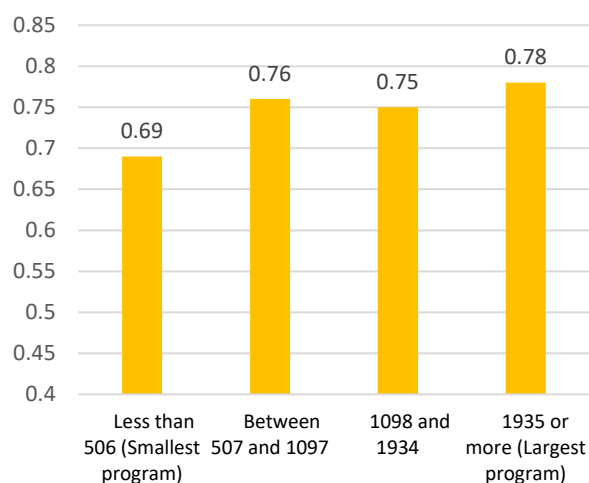
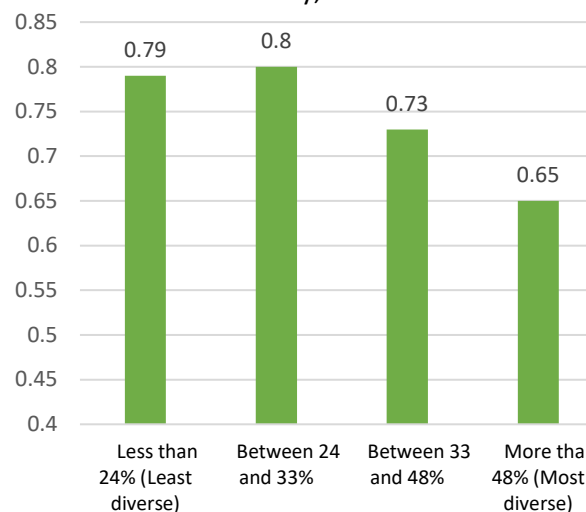


Chart 12. Pass Rate by Racial/Ethnic Diversity, 2018-2022

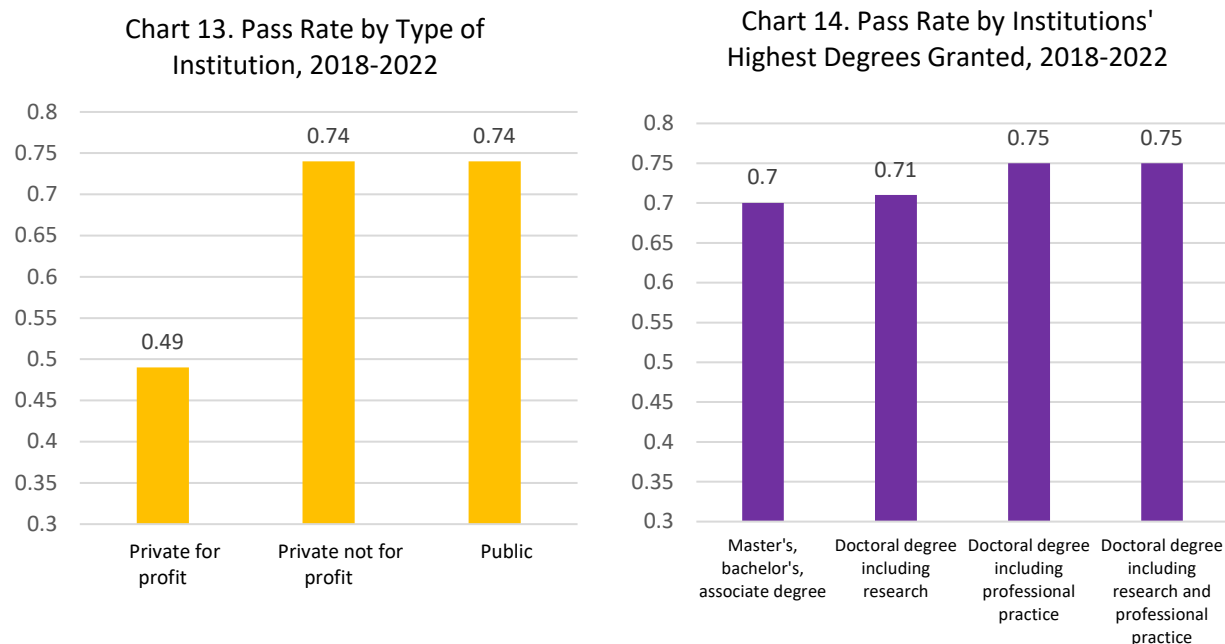


Findings presented in **Appendix Table A-1** suggest that a **higher share of examinees from historically marginalized groups earned their MSWs from institutions whose characteristics might have been negatively associated with Clinical exam outcomes.** For example, a higher share of Black examinees (31.48%) than white examinees (23.97%) obtained their MSWs from a program with the smallest number of Clinical examinees. At the same time, large percentages of Hispanic, Asian, and Black examinees earned their MSWs from institutions where nearly half of the examinees were from historically marginalized groups. Only about 15% of white examinees earned their MSWs from such an institution.

Type of Educational Institutions

Chart 13 shows that the type of educational institution that examinees attended was associated with exam pass rates. Examinees who attended private, for-profit institutions had a pass rate as low as 0.49. This low pass rate for examinees from private, for-profit institutions was consistently documented in the literature (Espahbodi et al., 2023; Mittelstaedt & Morris, 2017; Spector et al., 2020). The finding confirms that the same pass rate disparity by educational institution type exists in social work.

Chart 14 presents a related finding that examinees who attended doctorate-granting institutions had a higher pass rate than those who attended master's-granting institutions (0.75 versus 0.70).



How were these predictors of exam outcomes distributed across the examinees' racial/ethnic groups? According to **Appendix Table A-1**, only 0.56 % of all examinees had earned their MSWs from private, for-profit institutions (There were only two private, for-profit institutions that housed accredited MSW programs.) About 0.5% of white examinees, 1.17% of

Black examinees, and 0.80% of examinees from unknown race/ethnic groups were graduates of private, for-profit programs. In addition, approximately 70% of all examinees, across most racial/ethnic groups, earned their MSWs from doctorate-granting institutions. The race/ethnicity differences in the types of institutions were not as clear as those in other predictors of exam outcomes.

Institution's Location and Undergraduate Students' Socioeconomic Status

Chart 15 shows that the exam pass rates significantly differed by the urbanity of institutions' locations. Examinees from institutions located in mid-size or large-size urban areas passed the exam at a higher rate than those from institutions in rural or suburban areas (0.76 versus 0.71) (Angelo et al., 2021). As previous evidence reported, examinees who graduated from institutions with a higher share of students in poverty had a lower pass rate than their counterparts (Espahbodi et al., 2023). According to Chart 16, examinees who had attended institutions where at least half of the undergraduate students were Pell Grant recipients had a pass rate of 0.64, far lower than the pass rate of 0.80 for the examinees whose institutions had less than 20% of students on Pell Grant.

Chart 15. Pass Rate by Urbanity of Institution's Location, 2018-2022

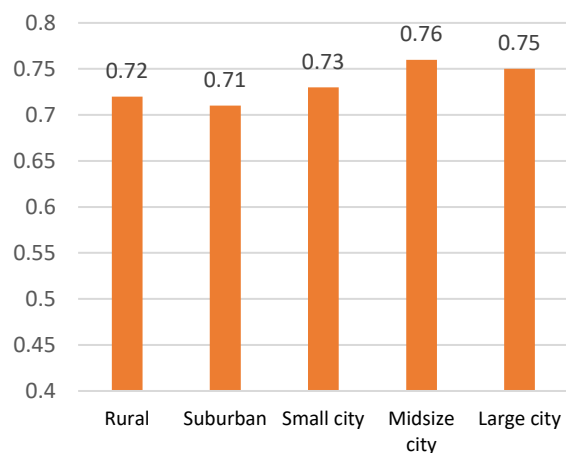
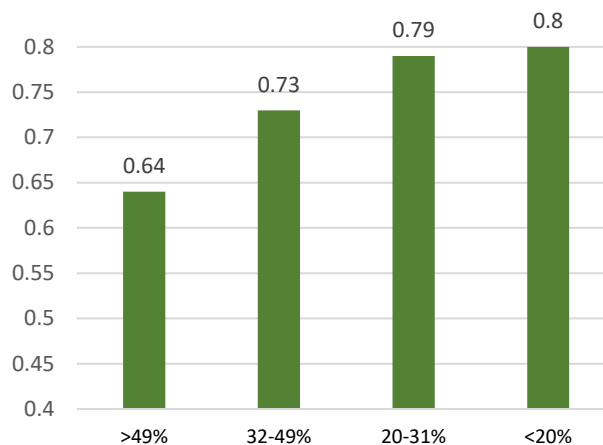


Chart 16. Pass Rate by Percent of Students on Pell Grant, 2018-2022

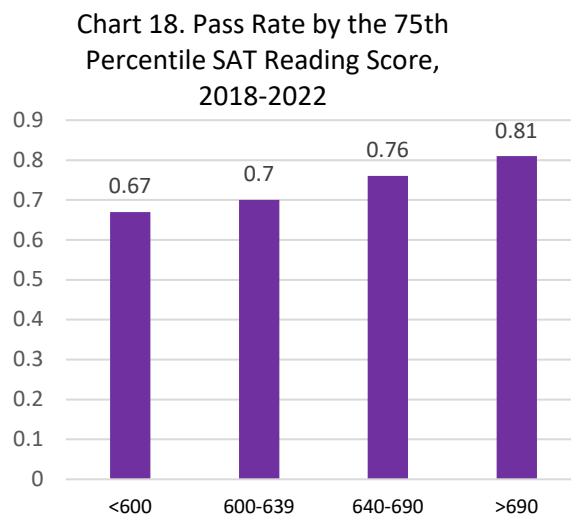
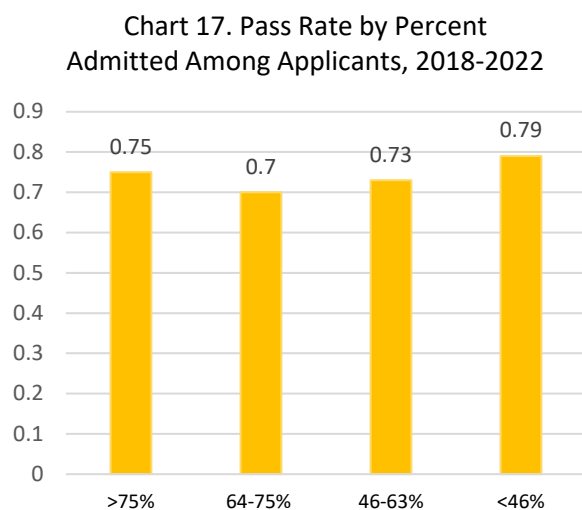


Appendix Table A-1 presents the locations of educational institutions by examinees' race/ethnicity. The finding suggests no significant patterns in the institutions' locations by racial/ethnic groups. However, **significantly greater shares of Black (35.91%) and Hispanic/Latino (36.26%) examinees, compared to 15.62% of white counterparts, had attended institutions where at least half of undergraduate students were Pell Grant recipients.**

Institution's Undergraduate Admission Selectivity

Chart 17 presents another well-established relationship in the literature between the selectivity of educational institutions and students' licensing exam outcomes (Chaparro, 2020; Espahbodi et al., 2023; Wightman, 1998). The chart demonstrates that examinees from more selective institutions (in terms of the percentage of applicants admitted) had a higher pass rate than examinees from less selective institutions. Examinees whose institutions accepted less than half of the undergraduate applicants had a pass rate of 0.79, but those whose institutions accepted somewhere between 67 and 79% of the applicants had a pass rate of 0.70.

Interestingly, **Appendix Table A-1** shows that nearly 41% of Asian and 34% of Hispanic/Latino examinees, compared to 22% of white examinees, had earned their MSWs from more selective institutions (i.e., with less than a 45% acceptance rate for undergraduate admission). White examinees had the lowest percentage of attendance at more selective institutions among all racial/ethnic groups.



As previous studies have reported (Chaparro, 2020; Espahbodi et al., 2023; Wightman, 1998), Charts 18, 19, and 20 show that success on the exam was associated with the admission selectivities of the educational institutions that examinees had attended. For example, the examinees who attended institutions where the 75th percentile SAT reading score for admitted first-year students was above 690 had a pass rate of 0.81, significantly higher than the pass rate of 0.67 for those from institutions where the score was below 600.

Charts 19 and 20 present similar pass rate gradients by undergraduate admission selectivity using the 75th percentile SAT math and ACT composite scores. Examinees from institutions where the 75th percentile ACT composite score was higher than 31 had a pass rate of 0.81. This was significantly higher than the pass rate of 0.64 for examinees who had attended institutions whose 75th percentile score was lower than 25.

Chart 19. Pass Rate by the 75th Percentile SAT Math Score, 2018-2022

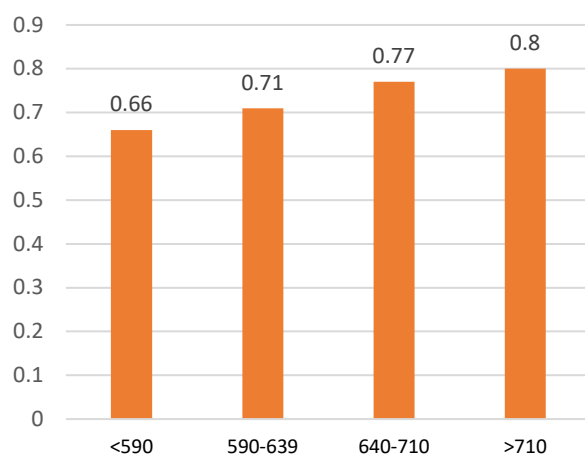
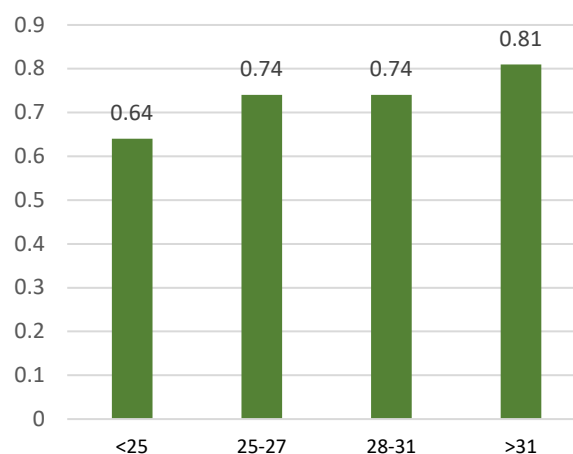


Chart 20. Pass Rate by the 75th Percentile ACT Score, 2018-2022



According to **Appendix Table A-1**, the highest shares of Black and Hispanic students were from institutions with the lowest 75th percentile SAT scores (both reading and math) and ACT composite scores. For example, while 16.22% of white examinees had attended institutions with below 600 for the 75th percentile SAT reading score, 23.85% of Black examinees and 23.12% of Hispanic/Latino examinees attended such institutions. These findings indicate that higher percentages of examinees from historically marginalized groups had attended institutions whose characteristics might have been negatively related to exam passage.

Pass Rate Disparities by Community Characteristics and Region

Consistent with the empirical studies of other licensing exams, success on the ASWB Clinical exam was significantly associated with community-level household incomes and racial income inequalities.

Examinees in high-income areas and areas with racial/ethnic income equality had a higher pass rate than those from less privileged areas. Pass rates also varied significantly by region of residence.

Household Income and Racial Income Inequality

As Chart 21 shows, examinees from low-income zip code areas had a significantly lower pass rate than examinees from high-income areas. Examinees in low-income areas (with annual

median household incomes less than \$57,665) had a pass rate of 0.63, but those in high-income areas (with incomes more than \$92,881) had a pass rate of 0.83.

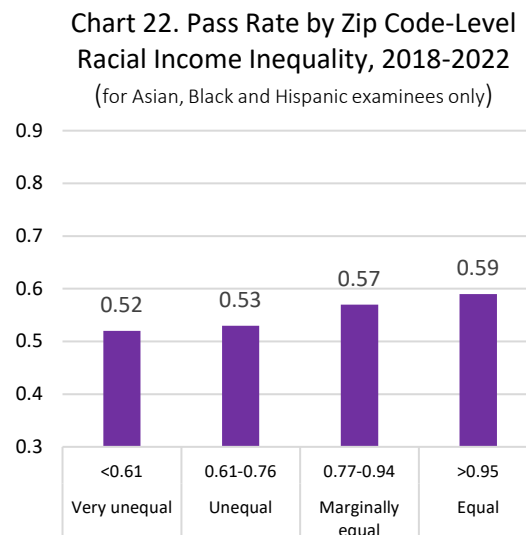
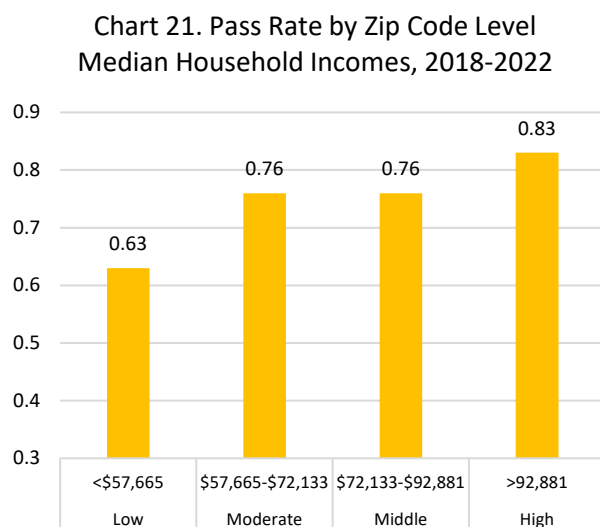


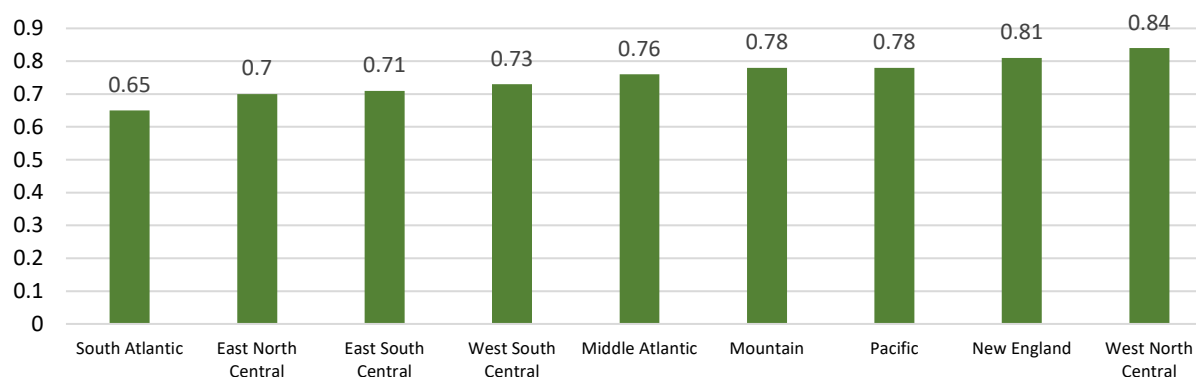
Chart 22 similarly demonstrates that examinees from areas with a high level of racial/ethnic income inequality were less likely to pass the exam than those from areas without such inequality. Asian, Black, and Hispanic/Latino examinees in zip code areas where their incomes were less than 61% of the incomes of white residents had a pass rate of 0.52. In the areas where the marginalized groups had at least 95% of the incomes of their white counterparts, the pass rate was 0.59. These findings were consistent with those reported for the CPA exam (Espahbodi et al., 2023).

Appendix Table A-1 presents an interesting finding about the zip code–level median annual household incomes for each racial/ethnic group. While nearly 47% of Asian and 29% of white examinees appeared to reside in zip code areas with an annual median income of at least \$93,000, only 6.6% of Black and 13.99% of Hispanic/Latino examinees did so. On the other hand, **more than 58% of Black examinees lived in the lowest income areas** (with an annual median household income below \$57,665). Only about 13% of Asian and 16% of white examinees lived in the lowest-income areas. About 26% of Hispanic/Latino examinees resided in the lowest income areas. Furthermore, Black examinees lived in areas with the highest level of racial income inequality. Appendix Table A-1 demonstrates that nearly 35% of Black examinees, compared to about 12% of Asian and 16% of Hispanic/Latino examinees, lived in zip code areas where Black residents' incomes were less than 61% of white residents' incomes. Only around 18% of Black examinees lived in areas where their incomes were at least about 95% of their white counterparts' incomes. These findings suggest that **a much higher share of Black examinees lived in economically segregated areas** than examinees from other historically marginalized groups. This means that higher percentages of Black examinees lived in communities with characteristics that may have had a negative effect on their exam performance.

Region of Residence

Chart 23 shows that Clinical exam pass rates significantly differed by examinees' region of residence. Examinees living in the South Atlantic region (DE, DC, FL, GA, MD, NC, SC, VA, WV) had the lowest pass rate of 0.65, much lower than the highest rate of 0.84 for those in the West North Central area (IA, KS, MN, MO, NE, SD, ND). Interestingly, there is a clear regional gradient in the pass rates, as shown in Chart 23. Although previous studies from other regulated professions (e.g., Falcone & Hamad, 2012) also presented a regional disparity in exam pass rates, the literature does not yet provide any theoretical explanations as to why such a regional disparity is observed.

Chart 23. Pass Rate by Region of Residence, 2018-2022



According to **Appendix Table A-1, racial/ethnic groups were not evenly distributed across the nine regions.** As high as 37% of Black examinees lived in the South Atlantic region, and nearly 46% of Asian and 42% of Hispanic/Latino examinees lived in the Pacific region. Although white examinees were most evenly distributed across the country, nearly 20% were from the East North Central area.

Net Effect of Race/Ethnicity on the Odds of Exam Failure

Logistic regression analyses suggested that if examinees from historically marginalized groups had the same demographic, educational, and employment characteristics and lived in similar institutional and community environments as white examinees, the Black–white disparity in the Clinical exam outcome could be reduced by about 20%, and the Hispanic/Latino–white disparity by around 28%.

Black examinees' exam outcomes were sensitive to institutional and community-level socioeconomic status and inequalities. On the other hand, Hispanic/Latino examinees' exam outcomes were explained more by their individual backgrounds.

As discussed earlier, a logistic regression analysis was conducted to understand the effects of race/ethnicity on the likelihood of failing first-attempt Clinical exams while taking the effects of demographic, institutional, and community characteristics into consideration. As shown in Table 2, the first race/ethnicity–only model (Model 1) included only race/ethnicity variables in the analyses. The second model (Model 2) added examinees' other demographic (other than race/ethnicity), educational, and employment characteristics to the first model. The last, full model (Model 3) added institutional and community-level variables to the second model. The goal was to observe how the odds ratios of the race/ethnicity variables declined between the models. **Appendix Table A-2** provides the full findings of the three regression models, including the odds ratios of variables other than race/ethnicity. Also, as discussed previously, the regression analyses did not include multiracial examinees or examinees whose race/ethnicity was unknown. Please note that due to high correlations between multiple variables (e.g., age groups and years of employment; SAT scores and percent of students on Pell Grant), the logistic regression models include only the necessary interrelated predictors of pass/fail outcome.

Table 2 presents the major findings in odds ratios of the race/ethnicity variables. The last two columns of the table show the percent reductions in the effect of race/ethnicity due to the added individual, institutional, and community-level variables. A comparison of the three regression models suggests that the effect of race/ethnicity remains large and significant, but the effect was reduced when other predictors were included in the model to explain exam failure, as discussed below.

Table 2. The Effects of Race/Ethnicity on First-Time Clinical Exam Failure, 2018-2022

	Model 1 (M1)		Model 2 (M2)		Model 3 (M3)		Percent (%) reduction in Odds Ratio between models	
	Race/Ethnicity only		Demographic, educational, and employment characteristics		Demographic, educational, employment, institutional, and community characteristics			
	O.R.	<i>p</i>	O.R.	<i>p</i>	O.R.	<i>p</i>	M1 vs. M2	M1 vs. M3
Asian	2.055	***	1.868	***	1.879	***	9.10	8.56
Black	6.483	***	6.019	***	5.193	***	7.16	19.90
Hispanic/Latino (White)	2.771	***	2.199	***	2.009	***	23.53	27.50

Notes: (1) O.R.: Odds Ratio; (2) *** $p < 0.001$; (3) Reference group is in parenthesis

According to the findings of the race/ethnicity-only model (Model 1), the odds of exam failure for Asian examinees were 2.055 times higher than the odds for white examinees. The odds of exam failure were 6.483 times higher for Black examinees and 2.771 times higher for Hispanic/Latino examinees than the odds for their white counterparts. In the individual characteristics model (Model 2), the odds ratios of all race/ethnic groups were reduced when examinees' demographic, educational, and employment characteristics were controlled for. For example, the odds of exam failure for Black examinees were **reduced by 7.16%** from 6.483 to 6.019 ($0.0716 = [6.483 - 6.019] / 6.483$), and the odds of exam failure for Hispanic/Latino examinees were **reduced by 23.53%** from 2.771 to 2.199 ($0.2353 = [2.771 - 2.119] / 2.771$).

When the full model (Model 3) added institutional and community-level variables to the logistic regression, **the odds of exam failure for Black examinees** — relative to the odds of white examinees' — **were reduced by 19.90%** from 6.483 to 5.193. **For Hispanic/Latino examinees, the odds of failure declined by 27.50%** from 2.771 to 2.009. **As suggested by the large reduction in the odds ratios between Model 1 and Model 3, Black examinees' exam outcomes appear to be sensitive to institutional and community-level socioeconomic status and inequalities.** This finding was closely in line with the finding reported by Espahbodi et al. (2023). On the other hand, **Hispanic/Latino examinees' exam failure was explained more by their individual backgrounds as indicated by the large reduction in the odds ratios between Model 1 and Model 2.** Asian examinees' odds of failure were reduced by about 9%, regardless of the inclusion of institutional and community-level determinants. This finding was not surprising because Asian examinees' overall characteristics were not as markedly different from white examinees as Black or Hispanic/Latino examinees' characteristics were.

Many additional regression analyses were conducted, although their findings were not presented in this report. For example, **when the same regression model was run with examinees in the South region only, the Black-white disparity in the exam failure rate was**

reduced by nearly 30% (from an odds ratio of 6.345 in the race/ethnicity–only model to 4.444 in the full model). This finding suggests that the included predictors had more explanatory power with a geographically homogenous sample of examinees. Likewise, when the regression analysis was conducted with examinees from private, nonprofit educational institutions, the included predictor variables reduced the Black–white disparity in the failure rate by 26%, from the odds of 6.415 in the base model to 4.747 in the full model.

Overall, the models explained relatively moderate levels of variance in the exam outcome, as shown in the pseudo- R^2 in **Appendix Table A-2**. It may be that variability in the exam data was high and that the full model did not include determinants that are critical to the exam outcome. As discussed earlier, many crucial predictor variables presented in Chart 1 were not available for this analysis. In addition, the institutional and zip code–level variables from the U.S. Census and NCES IPEDS data were, at best, only partial and crude proxies of the individual-level variables needed to test the conceptual framework depicted in Chart 1. The findings of this analysis should be interpreted with these limitations in mind.

Discussion

Summary

This study set out to shed light on the racial/ethnic disparities in the ASWB Clinical exam outcome. We tested the statistical significance of the individual, institutional, and community-level determinants of licensing exam outcomes identified in other licensed professions' literature with the ASWB Clinical exam data. The findings of our analyses are closely aligned with the extant literature. Older ages, gender, English use as a second language, beginning postsecondary education with an associate's degree, holding a BSW, delays in taking the licensing exam, and working in non-direct service positions were negatively associated with exam passage. Smaller MSW programs and institutions that largely served students from lower socioeconomic backgrounds were also negatively related to exam passage. Additionally, living in low-income communities was a significant predictor of exam failure. Our analyses showed that these negative predictors of exam outcomes were, in general, more prevalent among examinees from historically marginalized groups, indicating that race/ethnicity serves as a marker of socioeconomic disadvantages.

Our regression analyses tried to separate the net effect of race/ethnicity from the effects of correlated socioeconomic disadvantages in explaining exam outcomes. The findings showed that the effect of race/ethnicity declined considerably (by around 20–30% in the odds of failure) when the associated effects of socioeconomic disadvantages were controlled for. Nevertheless, race/ethnicity remained the most influential determinant of the ASWB Clinical exam outcome. These findings were not surprising given the limited predictor variables available in the ASWB exam data. As discussed earlier, our analyses could not incorporate the number of crucial predictors of exam outcomes identified in the conceptual framework. Despite the data

limitations, our overall findings on the determinants of the exam outcomes were consistent with the existing evidence in other professions' literature. The data limitations, however, should spur many further research questions, as discussed below.

Further Studies for Potential Interventions

The findings of this study may prompt many research questions. First and foremost, our analyses could not investigate the relationship between **examinees' academic backgrounds and performance** and their ASWB Clinical exam outcomes because we did not have access to such data. Some of the most comprehensive studies in bar and medical licensing exams revealed that racially disparate exam outcomes were primarily explained by examinees' GPAs in the academic programs and admission test scores (Rubright et al., 2019; Wightman, 1998). It would be important for the social work profession to replicate such studies to identify the factors contributing to the large racial disparities in the ASWB exam outcomes. Obtaining the examinees' academic backgrounds and performance data and linking them to the ASWB exam outcomes would require long-term collaboration between MSW programs and ASWB. Such partnerships may enable more comprehensive empirical research that can allow investigation of the relationship between the **characteristics of MSW programs** (e.g., faculty characteristics, curriculum, admission criteria, etc.) and ASWB exam outcomes. Because our analysis could not include any of these important predictors, further studies are necessary to unveil what was hidden in the relationships. **Collaboration between ASWB and MSW programs** will be critical in making the necessary data available for such studies.

Relatedly, collaborative research between ASWB and MSW programs can bring additional insights into program- and institution-level determinants of exam outcomes. Although social work candidates take Clinical exams many years after they graduate from MSW programs, the development of their clinical competence begins with MSW education and training. Our data analyses showed that some MSW programs did an excellent job of graduating Black and Hispanic/Latino candidates who would later succeed on the Clinical exam on their first attempts. A few examples of such MSW programs were at San Diego State University; the University of California, Berkeley; and the University of Texas at Austin. Black and Hispanic graduates of these programs passed the Clinical exams at around 80 to 90% on their first attempts. Given these exemplary programs, it may be fruitful to **investigate what sets those programs and their graduates apart from others**. Suppose there are common features in their student bodies, curricular contents, clinical training, or faculty qualifications that may be relevant to their graduates' later success on the Clinical exams. In that case, they may inform ideas for feasible and replicable interventions for other MSW programs or institutions.

Last, the findings of this study raised another question about **the labor market experiences of social work candidates, particularly Black candidates**. Little is known about how MSW graduates navigate the social work job market to secure employment in a clinical setting to gain the required postgraduate clinical supervision and if their experiences significantly differ by race/ethnicity. It is difficult to explain why a smaller share of Black MSW graduates held a direct service position and took many more years to attempt their first Clinical exams than

examinees from other racial/ethnic groups. It will be important to study if experiences of postgraduate clinical training are racially patterned and create barriers for Black candidates to develop clinical competence in a reasonable timeframe. States' regulatory rules and practices governing social work candidates' supervised training and ASWB exam registrations may also be an important topic for investigation. States vary in terms of requiring a specified amount of time to accrue supervision hours and take the ASWB Clinical exam. Therefore, it will be worthwhile to explore if there are any regulatory rules and practices that affect social work candidates' acquisition of the required postgraduate supervision.

Implications

Licensing exams play a critical role in verifying candidates' competence in a uniform and efficient way (Kane, 2005). For the public, a licensing exam certifies professional candidates' minimum competence to protect the public. For educational institutions, a licensing exam can help institutions externally validate student outcomes. For individuals, the assessment can provide useful feedback about areas for additional knowledge and skill development. As Kane (2005) explained, a high level of achievement on a licensing exam does not ensure success in practice, but a lack of adequate mastery of competencies may put clients at risk. Evidence from the legal and medical professions suggests significant relationships between exam outcomes and the indicators of public safety (Anderson & Muller, 2019; Cuddy et al., 2017; Tamblyn et al., 2007).

As with other professions' licensing exams, the ASWB exams follow strict test development standards, set by the American Psychological Association, the Joint Commission on Standards for Educational and Psychological Testing, the American Educational Research Association, and the National Council on Measurement in Education. Questions on the ASWB exams are reviewed for signs of potential bias at each step in the exam development process. Any questions identified as potentially biased, as well as those failing to accurately test candidates' knowledge, are not included on the exams. However, as with many licensed professions reviewed in the Exam Report Series, the social work profession has been challenged with racially disparate licensing exam outcomes.

The prevalence of racial disparities across many professions' licensing exams indicates that the causes of disparities are deeply rooted in the fabric of our socioeconomic systems. In response to the disparities in exam outcomes, many take a reductionist approach by blaming the exams or advocating to remove competence assessment in the licensure system. While no licensing exam may be perfect as an assessment tool for professional competence in the complexity of real practice environments (Kane, 2005), group differences in exam outcomes do not necessarily indicate that the exams are biased. They instead reflect persistent inequalities and segregation in our schools, communities, and workplaces that disproportionately and adversely affect people from low-income and historically marginalized backgrounds (Hauser & Heubert, 1998).

To maintain the integrity of the licensure system while not reinforcing and perpetuating the inequalities that continue to penalize people from historically marginalized groups, major professional stakeholders must come together to determine what further research is necessary and what interventions may be feasible and effective to address the problem. Doing this would be more challenging, yet more effective in narrowing the disparities than simply discarding the exam or the licensure system. When we can locate the sources of the disparities and know how to intervene strategically, we can reduce and even eliminate them in the long run, as they were socially created and not inherent in the demographic groups.

References

- Affrunti, N. W., & Rossen, E. (2023). *Examining racial-ethnic and gender differences on the Praxis School Psychologist Tests, September 2022-August 2023*.
<https://www.nasponline.org/research-and-policy/research-center/nasp-research-reports>
- Anderson IV, R., & Muller, D. T. (2019). The high cost of lowering the bar. *Georgetown Journal of Legal Ethics*, 32, 307. www.law.georgetown.edu/legal-ethics-journal/wp-content/uploads/sites/24/2019/08/GT-GJLE190055.pdf
- Angelo, B., Brasel, K., Stanfield, J., & Westfall, T. (2021). *Who are we missing? An empirical investigation of institution and program factors on graduate attempts on the CPA exam*.
https://nasba.org/wp-content/uploads/2021/08/Report-to-NASBA-on-CPA-Exam-Participation_08092021-Stanfield.pdf
- Association of Social Work Boards (2022). ASWB Policy Manual-v1. II. The Examinations.
<https://www.aswb.org/wp-content/uploads/2020/12/Section-II-The-Examinations-v1.2022.pdf>
- Chaparro, E. (2020). *Predictors for passing the Psychology License Exam* [Doctoral dissertation, Walden University]. <https://www.proquest.com/docview/2437388866?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations%20&%20Theses>
- Bline, D., Perreault, S., & Zheng, X. (2016). Do accounting faculty characteristics impact CPA exam performance? An investigation of nearly 700,000 examinations. *Issues in Accounting Education*, 31(3), 291–1230. <https://doi.org/10.2308/iace-51227>
- Cuddy, M. M., Young, A., Gelman, A., Swanson, D. B., Johnson, D. A., Dillon, G. F., & Clauser, B. E. (2017). Exploring the relationships between USMLE performance and disciplinary action in practice: A validity study of score inferences from a licensure examination. *Academic Medicine*, 92(12), 1780-1785.
http://www.stat.columbia.edu/~gelman/research/published/Exploring_the_Relationships_Between_USMLE.98203.pdf
- Eich, M. & O’Neill, T. (2007). *NCLEX delay pass rate study*. *NCLEX Psychometric Research Brief* (January 2007). National Council of State Boards of Nursing.
<https://www.ncsbn.org/public-files/delaystudy2006.pdf>
- Espahbodi, A., Espahbodi, L., Espahbodi, R., Walker, R., & White, G. T. (2023). Determinants of CPA exam performance. *Journal of Accounting Education*, 64, 100859.
<https://doi.org/10.1016/j.jacedu.2023.100859>
- Falcone, J. L. (2012). Compliance on the American Board of Pediatrics certifying examination and the importance of location and size on pass rates. *Clinical Pediatrics*, 51(5), 483-489.
<https://doi.org/10.1177/000992281243655>
- Hauser, R. M., & Heubert, J. P. (Eds.). (1998). *High stakes: Testing for tracking, promotion, and graduation*. National Academies Press.
<https://nap.nationalacademies.org/catalog/6336/high-stakes-testing-for-tracking-promotion-and-graduation>
- Kane, M. (2005). The role of licensure tests. *The Bar Examiners*, February, 2005.
<https://thebarexaminer.ncbex.org/wp-content/uploads/PDFs/740105-kane-1.pdf>

- Mittelstaedt, H., & Morris, M. (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. <https://doi.org/10.1080/08832323.2017.1313188>
- National Association of Social Workers/Association of Social Work Boards (2013). *Best practice standards in social work supervision*. <https://www.socialworkers.org/LinkClick.aspx?fileticket=GBrLbL4Buwl%3d&portalid=0>
- National Council on Teacher Quality (2019). *A fair chance: Simple steps to strengthen and diversify the teacher workforce*. https://www.nctq.org/dmsView/A_Fair_Chance
- Nayer, M. & Glover Takahashi, S. (2017). *What Ontario physiotherapist data says about risk to competence*. College of Physiotherapists of Ontario, Toronto, ON. https://www.collegept.org/docs/default-source/default-document-library/what-ontario-physiotherapist-data-says-about-risk-to-competence.pdf?sfvrsn=bb7cfa1_0
- Nettles, M. T., Scatton, L. H., Steinberg, J. H., & Tyler, L. L. (2011). Performance and pass rate differences of African American and white prospective teachers on PRAXISTM examinations: A joint project of the National Education Association (NEA) and Educational Testing Service (ETS). *ETS Research Report Series*, 2011(1), i–82. <https://doi.org/10.1002/j.2333-8504.2011.tb02244.x>
- Nguyen, J., Liu, A., McKenney, M., & Elkbuli, A. (2021). Predictive factors of first time pass rate on the American Board of Surgery Certification in General Surgery Exams: A systematic review. *Journal of Surgical Education*, 78(5), 1676–1691. <https://doi.org/10.1016/j.jsurg.2021.01.020>
- Rubright, J. D., Jodoin, M., & Barone, M. A. (2019). Examining demographics, prior academic performance, and United States Medical Licensing Examination scores. *Academic Medicine*, 94(3), 364-370. <https://doi.org/10.1097/ACM.0000000000002366>
- Sharpless, B.A. (2021). Pass rates on the Examination for Professional Practice in Psychology (EPPP) according to demographic variables: A partial replication. *Training and Education in Professional Psychology*, 15(1), 18–22. <https://doi.org/10.1037/tep0000301>
- Spector, N., Silvestre, J., Alexander, M., Martin, B., Hooper, J. I., Squires, A., & Ojemeni, M. (2020). NCSBN regulatory guidelines and evidence-based quality indicators for nursing education programs. *Journal of Nursing Regulation*, 11(2), S1-S64. [https://doi.org/10.1016/S2155-8256\(20\)30075-2](https://doi.org/10.1016/S2155-8256(20)30075-2)
- Tamblyn, R., Abrahamowicz, M., Dauphinee, D., Wenghofer, E., Jacques, A., Klass, D., ... & Hanley, J. A. (2007). Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. *Jama*, 298(9), 993-1001. <https://doi.org/10.1001/jama.298.9.993>
- Trinkle, B., Scheiner, J., Baldwin, A., & Krull, G. (2016). Gender and other determinants of CPA exam success: A survival analysis. *The Accounting Educators' Journal*, 26, 101–117. <https://www.aejournal.com/ojs/index.php/aej/article/view/337>
- Wightman, L. F. (1998). *LSAC National Longitudinal Bar Passage Study. The Law School Admission Council (LSAC) research report*. <https://eric.ed.gov/?id=ED469370>
- Yeo, H. L., Dolan, P. T., Mao, J., & Sosa, J. A. (2020). Association of demographic and program factors with American Board of Surgery qualifying and certifying examinations pass rates. *JAMA Surgery*, 155(1), 22–30. <https://doi.org/10.1001/jamasurg.2019.4081>

Appendix Tables

Table A-1. Percentage Distributions of First-Time Clinical Examinees' Individual, Institutional, and Community Characteristics (N=88,678)

	All	Asian	Black	Hispanic	Multi	Unknown	White
	100.00	3.55	14.13	11.92	5.32	1.55	63.54
Age group							
<=29	25.59	24.00	18.87	22.13	23.57	12.75	28.30
30–34	28.56	33.82	26.09	34.10	28.21	26.37	27.85
35–39	16.28	18.88	17.77	19.32	17.54	21.92	14.98
>=40	29.58	23.30	37.26	24.44	30.67	38.97	28.87
Male	12.94	14.88	12.00	13.64	13.73	18.94	12.70
English as a secondary language	7.12	32.77	3.63	33.36	9.44	6.26	1.36
Holding an associate's degree	21.10	15.61	24.03	30.34	25.58	22.21	18.61
BSW	33.64	24.73	40.47	32.09	31.08	27.17	33.27
Years since MSW							
Less than 1	3.81	2.16	1.72	2.36	4.52	2.99	4.60
1–2	16.70	10.81	14.64	11.38	15.44	12.38	18.70
3–4	38.65	41.32	30.93	42.43	37.87	36.56	39.63
5–6	18.00	20.22	20.48	22.12	20.03	20.10	16.32
More than 6	22.84	25.49	32.23	21.71	22.15	27.97	20.76
Years of employment							
<=1	13.82	10.33	15.79	11.27	14.91	12.31	14.01
2–3	40.63	40.08	29.15	39.12	38.23	37.22	43.77
4–5	22.39	24.92	23.22	26.09	24.52	22.65	21.19
6–9	12.98	14.27	17.22	14.83	13.22	15.88	11.53
>=10	10.18	10.39	14.62	8.69	9.12	11.94	9.50
Job position							
Direct service	62.30	61.92	53.95	57.32	61.09	57.68	65.33
Other	20.59	23.27	24.74	26.46	22.19	23.82	18.21
Not applicable	5.98	5.12	5.97	6.03	5.88	8.74	5.97
Administrative work	11.12	9.69	15.34	10.19	10.84	9.76	10.49
Size of graduate program							
Less than 506	25.68	25.05	31.48	28.25	25.90	23.60	23.97
Between 507 and 1097	28.70	29.78	23.97	29.97	29.53	29.50	29.37
1098 and 1934	23.76	14.94	24.61	14.36	19.88	21.34	26.21
1935 or more	21.86	30.23	19.93	27.42	24.69	25.56	20.45
Diversity of graduate student body							
Less than 24%	23.97	8.68	11.47	7.36	15.93	17.41	31.55
Between 24 and 33%	25.17	22.35	17.65	15.49	23.38	22.80	29.03
Between 33 and 48%	24.71	25.52	29.55	21.37	25.60	26.44	24.09
(More than 48%)	26.15	43.45	41.33	55.78	35.09	33.36	15.33

Type of institution							
Missing	1.55	2.73	1.36	0.81	1.46	1.97	1.66
Private, for-profit	0.56	0.19	1.17	0.28	0.36	0.80	0.51
Private, nonprofit	33.79	38.40	31.55	35.66	34.79	39.26	33.46
Public	64.10	58.68	65.92	63.24	63.39	57.98	64.37
Highest degree granted by the institution							
Missing	4.55	8.68	3.62	7.60	4.71	6.12	3.90
Master's, bachelor's, associate's	7.05	4.35	6.67	7.35	7.85	7.79	7.14
Doctoral, w/ research	11.43	6.68	13.96	9.74	11.33	10.85	11.48
Doctoral, w/professional practice	6.94	7.15	6.36	7.28	7.02	5.54	7.02
Doctoral, w/ research and professional practice	70.03	73.14	69.39	68.02	69.09	69.7	70.46
Urbanity of institution							
Missing	1.55	2.73	1.36	0.81	1.46	1.97	1.66
Town or rural	4.62	2.10	3.97	3.67	4.65	4.73	5.08
Suburban	19.40	13.83	19.15	19.31	18.65	19.52	19.84
Small city	14.75	8.04	13.35	8.88	11.39	10.71	16.91
Midsize city	13.22	10.33	15.14	9.36	13.39	12.24	13.69
Large city	46.45	62.97	47.02	57.96	50.47	50.84	42.80
Percent of students on Pell Grant							
Missing	3.56	3.91	4.14	1.95	3.78	3.50	3.69
(Less than 20%)	24.02	30.48	18.69	23.37	25.52	26.07	24.79
Between 20–31%	27.57	26.61	18.99	19.09	26.01	26.07	31.30
Between 32–49%	22.96	15.26	22.27	19.33	19.60	19.45	24.59
Greater than 49%	21.88	23.74	35.91	36.26	25.10	24.91	15.62
Admission selectivity							
Missing	7.77	6.23	9.51	5.92	8.70	6.92	7.76
>75% accepted	23.96	15.16	20.61	16.35	21.21	20.98	26.93
64–75% accepted	20.69	17.42	22.36	15.30	18.07	18.43	21.79
46–63% accepted	22.82	20.44	23.91	28.49	22.59	22.29	21.68
<46% accepted	24.76	40.75	23.62	33.94	29.42	31.39	21.84
75th PCTL SAT Reading score							
Missing	13.78	12.11	16.52	12.05	13.62	12.96	13.61
<600	18.17	14.24	23.85	23.12	18.10	17.84	16.22
600–639	20.00	20.53	19.31	22.18	20.64	20.32	19.65
640–690	23.44	14.69	21.76	15.98	19.9	20.83	26.05
>690	24.62	38.43	18.55	26.67	27.75	28.04	24.46
75th PCTL SAT Math score							
Missing	13.78	12.11	16.52	12.05	13.62	12.96	13.61
<590	16.40	12.02	22.35	21.00	16.76	15.95	14.44
590–639	20.33	20.41	18.40	21.04	19.50	20.17	20.69
640–710	23.18	15.48	22.49	17.18	20.00	20.32	25.22

>710	26.32	39.99	20.23	28.74	30.12	30.59	26.03
75th PCTL ACT score							
Missing	14.30	14.34	16.07	14.43	15.70	14.79	13.76
<25	14.06	13.22	22.87	21.34	15.19	13.91	10.68
25–27	25.20	21.61	20.72	23.57	22.42	22.72	27.00
28–31	18.26	10.11	18.38	11.53	15.76	17.26	20.19
>31	28.17	40.72	21.96	29.12	30.93	31.32	28.37
Zip code–level median household incomes by race							
Missing	10.18	3.02	5.28	2.70	-	-	3.37
< \$57,665	22.22	12.71	58.18	26.17	-	-	16.41
\$57,665–\$72,133	22.73	22.47	15.82	36.06	-	-	24.23
\$72,134–\$92,881	22.42	15.10	14.13	21.08	-	-	27.35
(> \$92,881)	22.45	46.69	6.60	13.99	-	-	28.63
Racial income ratio group							
Missing	-	3.02	5.28	2.70			
<0.61	-	12.30	34.63	16.37	-	-	-
0.61–0.76	-	6.68	26.12	28.31	-	-	-
0.76–0.95	-	26.51	16.10	32.33	-	-	-
> 0.95	-	54.51	17.88	20.29	-	-	-
Region of residence							
New England	8.18	4.55	4.28	4.61	6.62	9.47	10.02
Middle Atlantic	14.49	14.84	11.85	12.13	13.85	18.06	15.47
East North Central	17.14	10.52	16.33	8.50	12.86	12.38	19.79
West North Central	6.16	4.29	2.43	1.76	4.26	3.86	8.13
South Atlantic	17.68	9.47	36.99	12.53	17.42	17.99	14.83
East South Central	3.34	0.45	5.27	0.47	1.48	1.97	3.81
West South Central	6.76	4.90	10.09	8.56	6.22	5.75	5.86
Mountain	8.93	5.28	2.35	9.58	10.61	7.87	10.36
Pacific	17.31	45.71	10.42	41.86	26.69	22.65	11.74

Note: - denotes that data are not available.

Table A-2. Logistic Regression of First-Time Clinical Exam Failure

	Model 1		Model 2		Model 3	
	O.R.	<i>p</i>	O.R.	<i>p</i>	O.R.	<i>p</i>
Race						
Asian	2.055	***	1.868	***	1.879	***
Black	6.483	***	6.019	***	5.193	***
Hispanic (White)	2.771	***	2.199	***	2.009	***
Male			1.321	***	1.315	***
English as a secondary language			1.621	***	1.614	***
Holding an associate's degree			1.908	***	1.825	***
BSW			1.774	***	1.708	***
Years since MSW						
Less than 1			1.090		1.151	**
1–2 (3–4)			1.289	***	1.276	***
5–6			1.332	***	1.313	***
More than 6			1.616	***	1.582	***
Job position						
Other			2.006	***	1.963	***
Not applicable			2.118	***	2.104	***
Administrative work (Direct service position)			1.396	***	1.388	***
Size of graduate program						
Less than 506					1.159	***
Between 507 and 1097					1.057	+
1098 and 1934 (1935 or more)					1.220	***
Diversity of student body						
(Less than 24%)						
Between 24 and 33%					0.936	*
Between 33 and 48%					1.106	***
More than 48%					1.147	***
Percent of students on Pell Grant						
Missing					1.278	***
Greater than 51%					1.300	***
Between 39–51%					1.228	***
Between 27–38% (Less than 27%)					1.001	
Zip code–level median household incomes by race						
Missing					1.154	**
< \$57,665					1.252	***

\$57,665 and \$72,133					1.100	***
\$72,134 and \$92,881					1.153	***
(> \$92,881)						
Exam Year						
2018	1.0986	***	1.228	***	1.222	**
2019	1.0422		1.146	***	1.143	**
2020	0.9622		1.018		1.014	
2021	0.9844		1.002		1.003	
(2022)						
Model Statistics						
Cox-Snell R^2	0.0976		0.1465		0.1516	
Nagelkerke R^2	0.1436		0.2155		0.2230	
Likelihood Ratio Test: Chi-square (<i>df</i>)	8479 (7)	***	13082 (18)	***	13576 (32)	***
N	82591		82591		82591	

Note: The multiracial group and examinees with no race/ethnicity information were excluded from the logistic regression analysis. Reference groups are in parentheses.

+ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$