Asian American Discrimination in Harvard Admissions*

Peter Arcidiacono
Duke University
NBER & IZA

Josh Kinsler
University of Georgia

Tyler Ransom
University of Oklahoma
& IZA

April 16, 2021

Abstract

Using detailed admissions data made public in the *SFFA v. Harvard* case, we examine how Asian American applicants are treated relative to similarly situated white applicants. Our preferred model shows that typical Asian American applicants would see their average admit rate rise by around 1 percentage point (19%) if they were treated as white applicants. We show that one of the channels through which Asian Americans are penalized is the personal rating and that including the personal rating cuts the Asian American penalty by less than half. While identifying the causal impact of race using observational data is challenging because of the presence of unobservables, this concern is mitigated in our setting. There is limited scope for omitted variables to overturn the result because (*i*) Asian Americans are substantially stronger than whites on the observables associated with admissions and (*ii*) the richness of the data yields a model that predicts admissions extremely well.

*We are grateful to Peter Blair, Rafael Dix-Carneiro and Chunbei Wang for helpful comments. Peter Arcidiacono served as an expert witness for Students for Fair Admissions, Inc. (SFFA) in the *SFFA v. Harvard* case. SFFA is not funding his work on this paper. Josh Kinsler worked as a consultant for SFFA in the *SFFA v. Harvard* case. SFFA is not funding his work on this paper. The views expressed and conclusions reached in this paper are those of the authors; they do not purport to reflect the views of SFFA. To the extent this paper relies on records from the *SFFA v. Harvard* case, it relies solely on the public records from the case.

1 Introduction

For years, there has been a perception that Asian American applicants to elite US colleges and universities are held to a higher standard than similarly situated white applicants (Golden, 2006; Fuchs, 2019). Despite this public perception, empirical work on the topic is scarce, primarily due to lack of data. Universities tightly guard access to admissions data, and even the criteria by which universities score their applicants is often unknown. The *SFFA v. Harvard* case provided unprecedented access to Harvard's admissions process. Using information made public through this lawsuit, we show that Asian American applicants are penalized relative to their observationally-equivalent white counterparts.

The data we analyze covers six admissions cycles for applicants who, if they were to graduate in four years, would have done so as the Classes of 2014–2019. The Harvard admissions data is exceptionally rich. In addition to many demographic, geographic, and academic measures, the data include information on internal Harvard ratings that influence admissions decisions. These include Harvard admissions officers' ratings of the applicants overall as well as ratings on academics, extracurriculars, athletics, and personal qualities. It also includes the admissions officers' ratings of the letters submitted by high school counselors and teachers. Finally, the data include information on alumni interviews of the applicants in the form of an overall score and a personal score.

In this paper, we focus on typical Harvard applicants, or applicants who do not belong to one of the following groups: recruited athletes, legacies, donor connections, and children of faculty and staff (ALDC). These special applicants are predominantly white, receive large preferences in admissions, and are evaluated differently than typical applicants.[1] Our primary sample excludes these special groups in an effort to make reasonable comparisons across similarly situated applicants. Importantly, more than 97% of Asian American applicants are not ALDC, meaning that our sample covers the overwhelming majority of Asian American applicants. However, we also show that our key findings regarding Asian American discrimination are robust to the inclusion of ALDCs.

Among typical applicants, Asian Americans actually have a slightly higher unconditional

[1]See Arcidiacono, Kinsler, and Ransom (forthcoming) for details on the racial composition of ALDC applicants and how they are treated differently in Harvard's admissions process.

admit rate than whites. But as we show in Section 3, these unconditional admit rates mask substantial differences in qualifications between the two groups. While it is widely understood that Asian American applicants are academically stronger than whites, it is startling just how much stronger they are. During the period we analyze, there were 42% more white applicants than Asian American applicants overall. Yet, among those who were in the top ten percent of applicants based on grades and test scores, Asian American applicants outnumbered white applicants by more than 45%.[2]

Of course, Harvard values more than just academics. And here, too, Asian American applicants as a whole perform as well or better than white applicants on most of Harvard's ratings. But Harvard's ratings may also be affected by racial preferences and penalties. Indeed, Harvard acknowledges that race, in the form of preferences for under-represented minority groups (URMs), is one of the inputs into the overall rating (Day 4 Trial Transcript, p. 50).[3] Consistent with this, we find large racial gaps in the assignment of the overall rating conditional on academic strength. Similar patterns hold for the personal rating, suggesting that this measure is also directly influenced by race. Further, we show that racial groups who have observed characteristics associated with lower overall and personal ratings score higher on these ratings, again suggesting a direct role of race.

We estimate a model of Harvard admissions that aims to uncover the causal influence of Asian American status. Given our findings that race influences both the overall and personal ratings, our preferred model of admissions, described in Section 4, excludes both of these ratings. That said, we include numerous other variables that would capture differences in the non-academic attributes of the applicant pool. The set of controls available far outnumbers

[2]A similar pattern exists at Yale University. According to a litigation complaint (Document 1), there were 35% more white applicants as compared to Asian American applicants in 2017 and 2018. However, among the top ten percent of applicants based on grades and test scores, Asian American applicants outnumbered white applicants by 85%. Also similar to Harvard, white applicants in the top deciles of grades and test scores are admitted at significantly higher rates than their Asian American counterparts. For example, in the top decile of academic strength white applicants are admitted at a rate of 20.2%, while Asian American applicants are admitted at a rate of 14.3% (see p. 23).

[3]Past US Supreme Court rulings allow universities to enact affirmative action to increase the representation of URM applicants. For this reason, we focus primarily on comparisons between Asian American and white applicants, since neither of these groups should be impacted by affirmative action. However, we keep African American and Hispanic applicants in our sample since they help illustrate when and where racial preferences are active in Harvard's admission process. In Appendix E.2.1 we illustrate that estimating a model with just white and Asian American applicants also results in a statistically significant penalty against Asian Americans.

past work on admissions and, as a result, the model fits the data extremely well.[4]

Our preferred admissions model shows a substantial penalty against Asian American applicants relative to their white counterparts. The average marginal effect of being Asian American is -1 percentage point. Given that the overall admit rate for Asian Americans is around 5 percent, removing the penalty would increase their admissions chances by roughly 19%.[5] Even if the personal rating is included in the model, the statistical case for a penalty against Asian Americans remains, with the penalty declining by less than half.[6]

We identify the effect of race on admissions using a selection-on-observables approach. The concern with this methodology is that race can proxy for differences in unobserved factors not fully captured by applicant characteristics and Harvard ratings. To avoid concerns about omitted variable bias, researchers have often turned to audit and correspondence studies in other economic settings where racial discrimination is a concern. However, this approach is infeasible for college admissions in the United States because it would be impractical to generate fictitious applicants to the level of detail that admissions offices review applicants.[7] Alternatively, researchers have exploited differential patterns in outcomes across groups to address concerns about unobserved factors when estimating racial disparities in decision making.[8] This approach is typically employed in settings where limited data expands the

[4]Previous work on college admissions primarily uses third party data which contains basic controls such as race, gender, test scores, and athlete/legacy status. Examples include Hurwitz (2011), Long (2004), Espenshade, Chung, and Walling (2004), and Espenshade and Chung (2005). An exception is Bhattacharya, Kanaya, and Stevens (2017), who have access to admissions data for a selective UK university that includes not only test scores, but also interview and essay scores.

[5]Our finding of a 19% penalty is large, considering the magnitude of other penalties in the discrimination literature. As a common example, consider the gender wage gap. The unconditional gender gap in earnings for full-time workers in the United States is approximately 20%, with the gap narrowing to roughly 9% after adjusting for a battery of worker and job characteristics (Blau and Kahn, 2017).

[6]Furthermore, Harvard's Office of Institutional Research (OIR) estimated a number of admission models which all found a significant Asian American penalty (Trial Exhibit P009 and Trial Exhibit P028).

[7]In addition to test scores, high school GPA, and demographic characteristics, applicants must also submit lengthy personal statements, teacher recommendation letters, and counselor recommendation letters. Further, at the most elite schools the majority of applicants interview with alumni. The US admissions system contrasts with that of many other countries, where an entrance exam is the main determining factor of admission. For studies of college admissions in other countries, see Ding and Lehrer, 2007 and Jia and Li, 2017 (China); Ketel et al., 2016 (Netherlands); Kirkeboen, Leuven, and Mogstad, 2016 (Norway); and Bordon and Fu, 2015 and Zimmerman, 2014, 2019 (Chile).

[8]See Bhattacharya, Kanaya, and Stevens (2017) in the context of college admissions. This approach has been used more commonly in the criminal justice setting, including Knowles, Persico, and Todd (2001), Anwar and Fang (2006), and Arnold, Dobbie, and Yang (2018). These methods often rely on strong assumptions about the underlying preferences of the decision maker (Canay, Mogstad, and Mountjoy, 2020).

scope for unobserved factors to overturn the impact of race.

While neither of these approaches are available to us, concerns about omitted variable bias are significantly reduced in our setting. The richness of the applicant data yields a model of admissions outcomes that matches the data incredibly well. The Pseudo R^2 of our preferred model is equal to 0.56, a value well above what is considered an excellent fit (McFadden, 1979). For comparison purposes, Espenshade, Chung, and Walling (2004) estimate racial preferences in elite college admissions and obtain a Pseudo R^2 of around 0.2. In Section 4 we provide further insight regarding the fit of our model by mapping our Pseudo R^2 value to model accuracy and a more traditional R^2 measure.

Not only is there limited scope for omitted variable bias, the evidence suggests that, if anything, we are likely understating the direct penalty Asian Americans face in the admissions process. Asian American applicants are significantly stronger than white applicants on the observable characteristics—outside of race—that Harvard values when making admissions decisions. While there may be differences in unobservables as well, researchers typically assume that groups that are stronger on the observed characteristics are also likely to be strong on the unobserved characteristics.[9] However, in our case, the estimated effect of being Asian American is negative *despite* Asian Americans being positively selected on the academic *and* non-academic observables associated with admission. Thus, if we were to follow the literature and assume Asian American applicants are also stronger on the unobserved factors affecting admission, the actual penalty would be larger than our estimates indicate.

Since this paper relies on reports from the *SFFA v. Harvard* case, it is important to note that Harvard's expert reached a different conclusion on whether Asian Americans were discriminated against in Harvard's admissions process. In Appendix E, we outline the key differences in the analysis and why we believe our approach to be correct. But it is essential to point out that, taking all of Harvard's modeling choices and simply removing the personal rating shows a significant admissions penalty against Asian Americans.[10]

Accepting the Harvard defense provides a blueprint for how to discriminate, given that

[9]Indeed, that selection on observed characteristics moves in the same direction as selection on unobserved characteristics provides the motivation for the empirical approaches of Altonji, Elder, and Taber (2005), Krauth (2016), and Oster (2019).

[10]See Exhibit 21 of Document 419-141 and direct testimony on pages 8–10 of Day 14 Trial Transcript.

(*i*) *every* model of the personal rating shows a significant penalty against Asian Americans, (*ii*) Asian Americans are stronger on the observables associated with the personal rating, and (*iii*) Harvard never produced evidence linking the personal rating with later success. An institution can simply invent a rating that penalizes a specific group and have a significant portion of the final admissions, hiring, or promotion decision depend on this rating. The fact that Harvard needed to rely on such a defense provides corroborating support for the findings of this paper.

2 Data and Admissions Process

Our primary data source is applicant-level data from the Classes of 2014–2019 produced by Harvard in the *SFFA v. Harvard* lawsuit. However, due to court protections, we do not use this data directly. Rather, we rely on publicly available documents such as expert witness reports or internal admissions office memos that were made public as part of the lawsuit. Among these publicly available documents, we rely most heavily on the plaintiff's expert witness rebuttal report (Document 415-9).[11]

In all, the data used in Document 415-9 consist of 142,728 typical domestic applicants who have complete application data.[12] For each applicant, the data contain details on academics, demographics, geography, ALDC status, and a host of other variables (e.g. parental education, whether the applicant applied for financial aid, etc.). As mentioned in the introduction, we exclude ALDC applicants from our analysis, meaning we estimate racial preferences and penalties only for typical applicants. As Appendix A makes clear, it is possible to obtain consistent estimates for typical applicants while excluding ALDC applicants even when they compete for the same admissions slots.

There are four main ethnic groups that make up Harvard's application pool: whites, Asian Americans, African Americans, and Hispanics.[13] While there is significant variation

[11] For a complete list of legal documents we rely on in this paper, see Appendix Table F1.

[12] Domestic applicants are those who are either US citizens or permanent residents. There are certainly some applicants who attend high school in the US who are not domestic applicants. There are also some US citizens who attend high school overseas. Trial Exhibit DX 042 shows annual application, admission and matriculation rates by race/ethnicity for the Classes of 2000–2017. During the period of our data, internationals make up about 18% of the applicant pool and about 11% of the admitted class.

[13] Harvard also keeps track of Native Americans and Pacific Islanders, but these groups are too small to

within ethnic groups (e.g. Laotian American vs. Korean American), we aggregate to these main categories because the available documentation suggests that Harvard does the same (see Trial Exhibit DX 042 and Trial Exhibit P164). Around 8% do not report their race, with some evidence that this group is disproportionately white and Asian American (Document 415-8, p. 81). These applicants remain in the estimation sample and are treated as a separate racial group.

Harvard admissions officers read each application and assign three classes of ratings: (*i*) overall; (*ii*) the profile ratings, which include the academic, extracurricular, athletic, and personal ratings; and (*iii*) the school support ratings which include the counselor rating and typically two teacher ratings. In addition, most Harvard applicants are also interviewed in person by an alumnus or alumna who lives close to their high school. The alumni rate applicants on a number of qualities, but only the overall and personal rating are included in the database. Ratings typically take on values between 1 and 5, with 1 being the best. Harvard's internal admission ratings are included in the applicant-level data.

Trial Exhibit P001 contains a summary of the criteria by which ratings are assigned for the Class of 2018. Among the profile ratings, the criteria for evaluating academics and extracurriculars are straightforward and generally coincide with what one would expect. The personal rating, which is meant to capture personal qualities such as likeability, courage, and kindness, is the most vague.[14] The reading procedures instruct the reader to score a 1 if the applicant's personal qualities are "outstanding," a 2 if they are "very strong," a 3 if they are "generally positive," a 4 if they are "bland or somewhat negative or immature," a 5 for "questionable personal qualities," and a 6 for "worrisome personal qualities."

After applications have been read and assigned ratings for each category, a subset of the

be separately analyzed, so we classify them as "other." We also create a separate category for those who fail to report any race or ethnicity. We classify multiracial students in the same way Harvard does (Document 415-8, p. 92). Multiracial students who have any amount of African ancestry are coded as African American, those with any amount of Hispanic ethnicity are coded as Hispanic (but African-American Hispanics would be coded as African American), then those with any amount of Asian ethnicity are coded as Asian American (so that half-Asian half-white would be coded as Asian), and finally whites are those whose ethnicity is only white.

[14]Harvard revised their reader guidelines the summer before the trial, providing much clearer guidance on the scoring of the personal rating. In contrast to previous reader guidelines, the 2023 reader guidelines explicitly state that race should not be considered when assigning the personal rating. See Trial Exhibit P633.

applicants are passed on to an additional reader, called the "third reader."[15] Provisional admissions decisions are then made at a subcommittee level called a "docket." In March, final decisions are made at the full committee level with all admissions officers present.[16]

3 Descriptive Analysis

We now turn to the characteristics of typical (that is, non-ALDC) Asian American and white applicants.[17] We begin by looking at their family backgrounds, showing that, on average, Asian American applicants come from poorer families than white applicants. Despite this, Asian Americans substantially outperform their white counterparts on academics. We then examine how Asian Americans and whites are rated by Harvard admissions officers and how these ratings are correlated with academic preparation.

3.1 Demographics

The first panel of Table 1 presents demographic characteristics for typical white and Asian American applicants as well as by whether or not they were admitted. The overall admit rate of white applicants over this period is 4.89% which is slightly lower than the 5.13% admit rate for Asian American applicants.

Asian American applicants are 4.49 percentage points more likely to be labeled disadvantaged by Harvard readers off a base of 6.36 percent.[18] Those who are labeled disadvantaged are significantly more likely to be admitted, and this alone removes the difference in admission rates between white and Asian American applicants: for those who are disadvantaged, the admit rate for whites (Asian Americans) is 11.22% (10.33%); for those who are not, the admit rate for whites (Asian Americans) is 4.46% (4.49%). Asian Americans are also more

[15] Second readers are present when the admissions officer is new or if the case is especially difficult.

[16] For four of the six application cycles in our sample (2016–2019), Harvard offered an Early Action program to its applicants. Full-committee admissions decisions for Early Action applicants were made before the end of December. Possible admissions outcomes include admission, rejection, or deferral to the regular application pool. Harvard states that Early Action applications are reviewed in exactly the same manner as non-early-action applications; see `https://college.harvard.edu/admissions/apply/first-year-applicants`.

[17] See Arcidiacono, Kinsler, and Ransom (forthcoming) for characteristics of ALDC applicants.

[18] Trial Exhibit P001 instructs the Harvard reader to code the applicant as disadvantaged if "the applicant is from a very modest economic background."

likely to be first-generation college students and to have applied for a fee waiver. Both of these variables are also positively correlated with admission, though not as strongly as the disadvantaged status variable.

It may be surprising that Asian American applicants are disadvantaged relative to white applicants given that Asian Americans in the US have higher household incomes (US Census Bureau, 2020).[19] There are at least two explanations. The first is that children of low income Asian American families perform much better in the classroom than other racial groups.[20] This stronger performance in turn results in being more at risk of applying to Harvard. Second, well-off Asian Americans may be more aware of the perception of discrimination against Asian Americans in admissions and hence may be less likely to report their race. This would be consistent with college admissions consultants often advising Asian American students to appear less Asian (English, 2015).

3.2 Academics

The second panel of Table 1 shows measures of the academic preparation of white and Asian American applicants. And here it is striking how much stronger Asian American applicants are. Using white applicants as a base, Asian Americans on average score 0.3 standard deviations better on both the SAT1 math and SAT2 subject tests, around 0.05 standard deviations better on high school grades, and take over 1.5 more AP exams with an average score that is 0.09 points higher.[21] While the average white applicant scores at the 53rd percentile of the academic index distribution, the average Asian American applicant is at the 63rd percentile. The academic index is a weighted average of the applicant's scores on the SAT1, SAT2, and high school grade point average (or class rank). It is used by Ivy League institutions to ensure recruited athletes meet minimum academic standards.[22]

[19] Indeed, there are more disadvantaged Asian American applicants and admits than disadvantaged white applicants and admits despite whites making up a much larger share of the applicant pool.

[20] The recent controversy over admissions to Stuyvesant HS in New York City is a case in point (Wong, 2019). Admission to Stuyvesant and other elite high schools is based on a standardized test score. Stuyvesant's 2019 admitted class was over 65% Asian American and over 21% white, but less than 5% Hispanic or African American. Mayor Bill de Blasio proposed a minimum admissions quota for low-income students, but that was met with immediate legal opposition.

[21] AP exam scores are only available in a subset of the admission cycles.

[22] See Document 415-8 footnote 29 for a more detailed discussion of the academic index.

The only measure of academic preparation that whites perform comparably on is the SAT1 verbal.[23]

The differences in academic achievement between white and Asian American applicants becomes even more staggering when looking at deciles of the academic index. The first set of columns in Table 2 show the number of applicants in each decile for whites and Asian Americans. Overall, there are 42.5% more white applicants than Asian American applicants.[24] But in the top decile there are *45.6% more* Asian American applicants than white applicants. In the bottom five deciles, there are over two white applicants for every Asian American applicant. But in the top three deciles, there are 11% more Asian American applicants than white applicants.

The differences in representation across academic index deciles are only relevant if the academic index is correlated with admission. And it is. No white or Asian American typical applicants were admitted from the bottom decile in any of the six admissions cycles, and less than 10% of white and Asian American admits come from the bottom five deciles. In contrast, 73% of white and Asian American admits come from the top three deciles. Additionally, the admit rate increases monotonically with academic index decile for both whites and Asian Americans.

But as shown in Table 2, admission rates conditional on academic index decile are quite different between white and Asian American applicants. From the fourth decile to the tenth, white applicants are over 20% more likely to be admitted than their Asian American counterparts in the same decile. To illustrate, whites in the top (tenth) decile have an admit rate of 15.3% compared to an Asian American admit rate of 12.7%. And from the fifth decile to the ninth, Asian Americans are admitted at a rate similar to whites one decile lower.[25]

[23]There may be a concern that Asian American applicants have SAT scores that are inflated due to retaking. Indeed, Table 2 of Goodman, Gurantz, and Smith (2020) shows that Asian Americans in the population are 10–14 percentage points more likely to retake the SAT than other racial groups, holding fixed varying factors. But it is not clear whether this is true for Harvard, especially in light of Asian Americans who report their race being disadvantaged relative to their white counterparts. Further, there is no evidence of Harvard using this argument in defense of their admission practices.

[24]Summing the first two columns of Table 2 indicates that there are 57,451 whites and 40,308 Asian Americans.

[25]Similar racial gaps in admit rates conditional on academic index decile are observed at Yale (Document 1), suggesting that Harvard's behavior towards Asian Americans is not unique among elite institutions.

3.3 Harvard Ratings

It is of course the case that Harvard values much more than academics. Indeed, if the academic index were used to decide admissions for white and Asian American applicants (holding the admissions decisions of everyone else fixed), the number of Asian Americans admitted would increase by 828, a 40% increase.[26] So in order to rationalize the similarity in the overall admission rate between white and Asian American applicants, it must be the case that Asian Americans are substantially worse on other characteristics Harvard values, or they are being discriminated against, or some combination thereof.

We now examine how well Asian American and white applicants score on Harvard's ratings. We focus on how the share of applicants who receive a 2 or a 1 on each of the ratings differs by race, referring to a 2 or a 1 as a high score. A 2 is a natural cutoff since, for almost every rating and racial group, the median admit receives a 2 while the median reject receives a 3.[27]

For each of the ratings, Table 3 shows the share of typical Asian American and white applicants who receive a high score.[28] Given the previous discussion, it is not surprising that Asian Americans score substantially better on the academic rating: 60.2% of Asian Americans receive a high score, compared to 45.3% of white applicants.

But Asian Americans also score well on many of the other ratings. Asian Americans are more likely to have high scores on the extracurricular rating and alumni overall rating and slightly more likely on both teacher ratings, the alumni personal rating, and the overall rating. They are slightly less likely to have high scores on the counselor rating, but the difference is less than 0.2 percentage points.

There are, however, two ratings where Asian Americans score significantly worse: the athletic rating and the personal rating.[29] Receiving an athletic rating of 2 does boost one's

[26] This number is obtained by fixing the total number of white and Asian American typical admits at their current level, 4884, and then randomly sampling from the tenth decile to fill the class. If we further added in the other racial groups, the number of Asian Americans would more than double, with Asian Americans making up more than 50% of admits.

[27] See Trial Exhibit P621. The exceptions are as follows: athletic rating (median admit and median reject both receive a 3); the alumni personal rating for all groups and the academic rating for Asian Americans (median admit and median reject both receive a 2); and the overall and academic ratings for African Americans (median admit receives a 2 but median reject receives a 4).

[28] We treat those who are missing a rating as not having received a high rating.

[29] When the Office of Civil Rights investigated Harvard, it was these two ratings that were found to be the

chances of admission. However, the share of typical applicants who receive a high score on the athletic rating is smaller than the corresponding share for any other rating. Additionally, a 4 on the personal or academic ratings virtually guarantees rejection for typical applicants. But there is no difference in the admit rate between those who get a 3 on the athletic rating versus those who get a 4. For typical applicants, then, the athletic rating appears to be less important than the other profile ratings in determining admissions outcomes.[30] Turning to the personal rating, not only do Asian Americans score worse than whites, they score worse than African Americans, Hispanics, and those not in one of the four major race/ethnic groups.[31] And the personal rating is strongly correlated with admission: 84% of white admits scored a 2 or better on the personal rating, compared to 18% of white rejects.[32]

The clear discrepancy between the personal rating and all other ratings is further shown by Harvard's OIR in Appendix Figure F2 for the Classes of 2007–2016. Of the characteristics that OIR analyzed, the personal rating is a clear outlier: Asian Americans are as strong or significantly stronger than whites on every rating except for the personal rating. Yet on the personal rating, they score more than 0.1 standard deviations lower—a fact that OIR could not statistically explain. The athletic rating is not included in this chart because legacies and athletes are excluded from their analysis and the primary purpose of the athletic rating is to distinguish recruited athletes.

3.4 Academics and Harvard Ratings

Given that Asian Americans are so much stronger on academics, could it be that excelling at academics comes at the cost of being appealing to Harvard on the personal front? This turns out to not be the case. In fact, the academic index is positively correlated with each of Harvard's ratings with the exception of the athletic rating.[33] Appendix Tables F2 through F4 show the share of applicants receiving a 2 or better for each of the ratings by academic

most subjective. See Trial Exhibit P555.

[30] Arcidiacono, Kinsler, and Ransom (forthcoming) show that, among non-recruited athletes, white LDC applicants score the best on the athletic rating. The high scores may be in part due to the sports Harvard offers, such as sailing.

[31] See the "Average" row in Table 5.6R of Document 415-9.

[32] See Table 4.1R of Document 415-9.

[33] There is no information in the public record about how the athletic rating is correlated with other factors for non-recruited athletes.

index decile and race. For ratings like the academic rating, virtually no one receives a 2 if they are in the bottom decile, and virtually everyone receives a 2 if they are in the top decile, regardless of their race. But for every rating and for every racial group, higher academic index deciles are associated with higher probabilities of receiving a 2 or better.

Figure 1 illustrates this pattern graphically for the academic, extracurricular, personal, and overall ratings. The share receiving a 2 or better increases significantly from the 1st to the 10th decile of the academic index for each rating and each racial group. For the academic and extracurricular ratings, the racial gaps in the share receiving a 2 or better within each academic index decile are fairly small. In contrast, the personal and overall ratings indicate large and consistent racial gaps in the likelihood of receiving a 2 or better within each academic index decile. Moreover, the ordering of the racial categories within an academic index decile is identical between the overall and personal rating, with African Americans receiving the highest share followed by Hispanics, whites, and Asian Americans, respectively. Note that when readers assign an overall rating, they are allowed to incorporate any factors deemed valuable to Harvard, including race.[34] Thus, the patterns observed for the overall rating are likely a reflection of racial preferences. Because the pattern for the personal rating mirrors that of the overall rating, this is suggestive evidence that racial preferences play a role in the personal rating as well.

Further evidence that racial preferences impact the personal rating is presented in Table 4. This table focuses on applicants in the top academic index decile and compares the probability of receiving a two or higher on each of Harvard's ratings for Asian Americans relative to the three other major racial/ethnic groups.[35] The first column of Table 4 shows the share of Asian Americans in the top academic index decile who receive a high score on each of the ratings. The second column shows how much higher (or lower) the similar share was for white students in the top decile, with the third column showing the corresponding percentage increase (or decrease) relative to Asian Americans. The results are sorted by the difference between the white and Asian American share. In the top decile, the share of whites receiving a high score on the personal rating was over seven percentage points higher

[34] See Document 421-9, pp. 259, 288 and 422.

[35] Over 44% of Asian American admits are in the top academic index decile.

than Asian Americans, a 33% increase. While whites in the tenth decile scored higher on other ratings as well, the gaps are substantially smaller.

The fourth column shows the lowest decile in the academic index distribution where whites would still have a higher probability of receiving a high score. For the personal rating, whites in the 6th academic index decile have a higher probability of receiving a high score than Asian Americans in the top academic index decile. For all of the other ratings, Asian American applicants in the top decile have a higher probability of receiving a high score or better than whites below the 9th decile.

While the Asian American and white comparisons illustrate that the personal rating stands out relative to the other ratings, the especially striking comparisons are with African American applicants. African American applicants in the top academic index decile are over twice as likely (111.57%) to receive a high score on the personal rating as their Asian American counterparts (47% for African Americans versus 22% for Asian Americans).

It is important to note that higher academic index deciles are associated with higher personal ratings for all racial groups. African American applicants in the top (10th) decile are also twice as likely to receive a high score as African American applicants in the third decile.[36] Yet, even African Americans in the *third* academic index decile are more likely to receive a high score on the personal rating than Asian Americans in the *tenth* decile.

The patterns in the personal rating mirror what we see in places where Harvard acknowledges race places a role: in the overall rating and in the admit rates themselves. The bottom panel of Table 4 shows the overall rating and admit rate for those in the top academic index decile. Like the personal rating, Asian Americans are rated the lowest and this is especially so when compared to African Americans and Hispanics.

The descriptive analysis strongly suggests that race plays a role in the personal rating in addition to the overall rating and admissions. It may affect the other ratings as well, but the personal rating is where the patterns are most stark. Based in part on these findings, our preferred admissions model discussed in the next section excludes the personal rating since its inclusion will tend to understate the role of race of admissions, though we return to this issue in Section 5.1.

[36] See Appendix Table F3.

4 A Model of Harvard Admissions

The descriptive evidence related to admissions suggests that there is scope for an Asian American penalty in the Harvard admissions process. We now turn to estimating a model of Harvard's admissions decisions, focusing in particular on measuring how being Asian American influences one's admissions outcome. In the sections below, we present our preferred admissions model and discuss the estimated Asian American penalty.

4.1 Admissions Model

The admissions data made available as part of the SFFA lawsuit cover six admissions cycles and include hundreds of variables describing each applicant.[37] It is not feasible to include every variable in every year since there would be as many regressors as admits. In the paragraphs that follow, we briefly discuss some of the key modeling choices that allow us to capture admissions decisions and the role of race in a simple, yet accurate manner.

The first key modeling decision is to pool the admissions data across years and estimate a logit model with indicators for admissions cycle.[38] The advantage of pooling the data is greater statistical power for uncovering some of the intricate patterns in admissions choices that are time-invariant. The drawback of the pooled model is that the relative importance of various applicant attributes may change over time. For example, if Harvard seeks to balance intended majors within each admissions cycle, then intended humanities majors are more valuable in years when there are relatively few. The pooled model can accommodate this variation through interactions between intended major and year. The question is which applicant characteristics are likely to have time-varying impacts. Fortunately, during the weeks and months that Harvard is making final admissions decisions, the admissions office publishes statistics about the makeup of the current admitted class, as well as how these numbers compare to previous classes. Admissions officers can use these "one-pagers" to generate similarly constituted admit classes over time, even if the applicant pool is changing.

[37] A similar discussion of our modeling approach is presented in Section 4 and Appendix C of Arcidiacono, Kinsler, and Ransom (forthcoming).

[38] The indicators for admissions cycle insure that in each year the average probability of admission matches that of the data.

We use these "one-pagers" as guidance and include in our pooled regression interactions of admissions cycle with applicant characteristics included in the "one-pagers" such as gender, docket, intended major, and disadvantaged status.[39]

In addition to indicators for applicant race and the above interactions, we incorporate a broad set of controls, including numerous measures of socioeconomic status, neighborhood and high school attributes, and academic aptitude, among others. We also control for many of Harvard's internal ratings, including the academic, extracurricular, athletic, the school support measures, and the alumni interviewer ratings. For each rating, we create separate indicator variables for rating levels from 1 to 5. We do not include either the overall rating or the personal rating. The overall rating is specifically designed to incorporate admissions preferences, including racial preferences, and is therefore an inappropriate control. Similarly, the personal rating is excluded since, as we showed in the previous section, it also appears to be influenced by racial preferences. In Section 5.1, we discuss additional evidence that the personal rating is influenced by race, but also show how the estimated Asian American admissions penalty is impacted by its inclusion in the model.

To allow for the possibility that racial preferences operate differently according to applicant disadvantaged status and gender, we interact each of these indicator variables with race. Arcidiacono (2005) shows that racial preferences for African Americans in admissions and financial aid vary with whether the applicant is low income. Additionally, African American applicants are disproportionately female (60%), so if Harvard is interested in gender balance within race, African American men may see larger preferences than African American women. This is in contrast to the applicant pool as a whole, which is less than 50% female.[40] We also interact race with indicators for early application status and missing SAT II average. The latter allows for differences in the distribution of missing scores by race since the observed test score distributions differ by race.

Our preferred model includes 128,422 applicants across the six admission cycles. This sample is smaller than the sample of typical applicants discussed previously for three reasons. First, neighborhood and high school characteristics are only available for domestic

[39]Section E.2.2 in the appendix illustrates that similar models estimated year-by-year result in an estimated Asian American penalty nearly identical to what we find using our pooled specification.

[40]See Document 415-9, Table B.3.2R.

applicants applying from within the US. Second, there is a small number of applicants who lack both teacher ratings. Finally, applicants whose characteristics perfectly predict rejection are excluded. The full set of controls for our preferred model is listed in Appendix B.[41]

4.2 Estimates of the Asian American Admissions Penalty

A subset of the estimated parameters for models with different sets of controls is displayed in Table 5. Column 1 controls for race and a handful of other demographic characteristics, yielding a coefficient on Asian American that is positive, small, and insignificant. This is consistent with the raw admit rates being similar for Asian American and white applicants. However, when academic characteristics are added, the Asian American coefficient becomes large and negative. This is consistent with Asian American applicants being much stronger in academics and Harvard putting significant weight on academics in the admissions decision. Adding further controls maintains the significantly negative coefficient on Asian American.

Column 5 shows estimates of our preferred model.[42] The coefficient on the Asian American indicator is negative and statistically significant (-0.466). Because the model contains interactions between race and gender and race and disadvantaged status, the results imply that—all else equal—male, non-disadvantaged Asian American applicants are penalized relative to similarly situated white applicants.[43] Although we do not display the coefficients in the table, the estimated Asian American penalty does not vary significantly with disadvantaged status, but it does vary by gender. The estimated coefficient on the interaction between Asian American and female is positive and significant (0.229), indicating that the Asian American penalty is smaller for female applicants.

The other coefficients in the estimated admissions model are consistent with Harvard's reader guidelines for evaluating applicants as well as their stated preferences for underrepre-

[41] Our preferred admissions model includes approximately 350 applicant characteristics. One might be concerned that the estimated penalty against Asian American applicants results in part from the model being overfit. However, Table 5 shows that a series of intermediate models with fewer controls indicate a negative and statistically significant penalty again Asian American applicants. The only intermediate admissions model that does not show an Asian American penalty is an overly sparse model that only controls for demographics and geography—meaning no academic or Harvard ratings variables are included.

[42] Column 6 illustrates how the coefficients change when we add the personal rating. This is discussed further in the next section.

[43] The model also includes an interaction between race and early applicant status, but these coefficients are close to zero and statistically insignificant.

sented minority groups and disadvantaged students. Applicants who receive a two or better on the academic or extracurricular rating are significantly more likely to be accepted.[44] Applicants who receive a four on either of these criteria see their chances of admission diminished relative to receiving a three.[45] The coefficients associated with being African American, Hispanic, or disadvantaged are all large, positive, and statistically significant.[46] Finally, conditional on all other observed attributes, early action applicants are more likely to be admitted than their regular admissions counterparts.

While the parameter estimates from our preferred specification indicate the existence of an Asian American penalty, it is difficult to understand the magnitude of the penalty from the coefficients alone. To put the magnitude of the Asian American penalty in context, we pursue two strategies. First, using the estimated coefficients, we show how the probability of admission would change for Asian American applicants if they had been treated as white applicants. Consider, for example, a male, non-disadvantaged Asian American applicant with a baseline probability of admission of p. The index of observables, Z, for this applicant according to the log odds formula is given by

$$Z = \ln\left(\frac{p}{1-p}\right) \tag{1}$$

which is the inverse of the standard logit formula. If this applicant were instead white, we would simply subtract the Asian American coefficient (-0.466) from the index so that the new admissions index would be $Z-(-0.466)$. The new admissions probability would then be given by $\frac{\exp(Z+0.466)}{1+\exp(Z+0.466)}$. A similar calculation can be made for various combinations of gender and disadvantaged status. The additional complication is that coefficients related to the interactions between Asian American and gender and Asian American and disadvantaged also need to be differenced out when applicable.

Table 6 lists the the results of these transformation exercises. The first entry in the

[44]Though not shown, a similar pattern holds for the athletic and personal ratings.

[45]Obtaining a rating of 5 on the extracurricular rating is an indication of substantial activity outside conventional extracurricular participation such as family commitments or term-time work. See the 2018 reader guidelines (Trial Exhibit P001).

[46]The interactions of African American and Hispanic with disadvantaged status are negative and statistically significant but the overall racial preferences for disadvantaged African Americans and Hispanics are still large.

table indicates that a non-disadvantaged, male, Asian American applicant with a baseline probability of admission of 1.0% would be admitted at a rate of 1.58% if treated as a similarly situated white applicant. This change reflects a 58% increase in the likelihood of admission. For other combinations of gender and disadvantaged status, the Asian American penalty is smaller.[47] As the baseline probability of admission increases, the percentage point increases are larger, but the percent increases are smaller: when the baseline admit rate is 25%, a non-disadvantaged, male, Asian American applicant would be admitted at a rate of 34.69% if treated as a similarly situated white applicant (a 39% increase).

The transformation exercises indicate that there is significant heterogeneity in the Asian American penalty according to gender, disadvantaged status, and the broader strength of the applicant. The average penalty faced by Asian American applicants will depend on the distribution of these characteristics in the applicant pool. Overall, the average marginal effect associated with Asian American is -1.02 percentage points, or a penalty of 19% off the base admit rate of 5.19%.[48]

In Table 7 we report the average marginal effect of being an Asian American applicant by admissions index decile. The admissions index deciles are created by ranking Asian American applicants according to their observable indexes; that is, taking the controls and multiplying by the coefficients of the preferred model. Because of how competitive Harvard is and how well the model fits the data, the average marginal effects are highly skewed. Those in the bottom 50% according to the admissions index have essentially no chance of being admitted, a reflection of how well the model fits the data. Hence the marginal effect for the bottom 50% of Asian American applicants is quite small at -0.02 percentage points, still a substantial penalty given their base admit rate of 0.04%. The average marginal effect rises with each decile with those in the top decile seeing an average marginal effect of 6.19 percentage points, or a penalty of 14%.

[47] For example, a disadvantaged, female, Asian American applicant with a baseline probability of admission equal to 1.0% would be admitted at a rate of 1.1% if treated as a similarly situated white applicant.

[48] See Table 8.2 of Document 415-9. Note that applicants who the model can predict perfectly are not included in the calculation. When these applicants are included, the average marginal effect falls slightly to -0.99%. We prefer the results excluding perfect predictions as race is only relevant for applicants who have a chance of admission. Either way of calculating the marginal effect results in a statistically significant effect at the 5% level.

5 Robustness of the Asian American Penalty

Asian American applicants to Harvard face a substantial admissions penalty relative to similarly situated white applicants. In this section, we demonstrate that our finding is robust to the inclusion of the personal rating and ALDC applicants. We also investigate the likelihood that our estimates could be driven by omitted variable bias.

5.1 Personal Rating

A key set of controls in the estimated admissions model are the internal Harvard ratings of applicants. However, we exclude two ratings from our preferred model, the overall and personal ratings. Harvard readers are explicitly allowed to incorporate an applicant's race in the overall rating, and, as a result, it is an improper control if the purpose of the model is to estimate racial preferences. In Section 3, we present descriptive evidence that the personal rating is also significantly influenced by applicant race.

In Appendix C we provide additional evidence that the personal rating is influenced by race, and is thus an improper control in an admissions model aimed at estimating racial preferences. We estimate a series of ordered logit regressions where the outcome is a rating of interest, say the extracurricular or personal rating. Our preferred ratings models are very similar to our preferred admissions models in the types of applicant characteristics included. Importantly, in each of the ratings models, we condition on all of the other Harvard ratings, excluding the personal and overall ratings.[49] While applicant race plays a statistically significant role in a handful of Harvard's ratings, race is most prominent in the personal and overall ratings. Further, in these two ratings, adding controls moves the race coefficients away from zero; the models for the other ratings generally show the coefficients moving towards zero as more controls are added. See Appendix Figure F1.

After controlling for hundreds of applicant characteristics, including Harvard's other ratings and measures of socioeconomic status, Asian Americans receive significantly lower per-

[49] We also estimate models without these ratings with similar qualitative findings. Note that, were Asian Americans to be penalized in other ratings besides the overall and personal ratings, controlling for these ratings would result in any Asian American penalties in the ratings to appear smaller than they actually are.

sonal and overall ratings. This occurs despite Asian Americans being stronger on the observables associated with these ratings; and indeed stronger on the observables associated with every modeled rating.[50] In contrast, African American and Hispanic applicants receive significantly higher personal and overall ratings. For a sense of the magnitudes, Asian Americans would see 20% higher odds of receiving a 2 or better on the personal rating if they were treated as white applicants. These odds would almost double if they were treated as African American applicants. Moreover, Asian American (African American) applicants are stronger (weaker) on the observed characteristics associated with high personal and overall ratings, making it unlikely that they are significantly weaker (stronger) on the unobserved factors impacting the personal rating.[51]

There are additional patterns in the assignment of the personal rating that suggest it is a tool to implement Harvard's preferences over the composition of their admits. For example, in the personal rating model the interactions between African American and female and African American and disadvantaged are significantly negative, implying racial preferences are muted for these two groups. The share of applicants who are female or disadvantaged is significantly higher for African Americans than for any of the other three major racial/ethnic groups, so if Harvard is interested in balancing within-race characteristics then we would expect to see muted preferences for African American applicants who were female or disadvantaged.[52] The only other rating that has this pattern is the overall rating, a rating that we know Harvard uses to directly implement preferences.[53]

Despite the preponderance of evidence indicating that the personal rating incorporates racial preferences, we also estimate the Asian American penalty in admissions when the personal rating is included. These results are shown in the last column of Table 5. We

[50] See Table B.6.11R in Document 415-9. The athletic rating was not modeled.

[51] In further support of this point, admissions data from the University of North Carolina-Chapel Hill (UNC) contradicts Harvard's claim that Asian American applicants have worse personal qualities (see Tables A.5.1 and A.5.2 of Document 160-1). Estimating a similar personal rating model using either in-state or out-of-state applicants to UNC reveals no differences in personal ratings for Asian American and white applicants. Note that the out-of-state admissions process to UNC is highly competitive, with an admit rate of only 13.5% (see Table 2.1 of Document 160-1)

[52] Table 3.1R of Document 415-9 shows descriptive statistics by racial/ethnic group, including share female and share disadvantaged.

[53] Section E.1 provides an extended discussion of the evidence showing that the personal rating incorporates racial preferences.

do this for two reasons. First, it allows us to determine how much of the Asian American penalty operates through the personal rating. Second, it provides a plausible lower bound for the true Asian American penalty. Adding the personal rating to our admissions model leaves a statistically significant penalty of -0.54 percentage points which, given the admit rate for Asian American applicants as a group, implies that Asian Americans would be 10% more likely to be admitted if treated as similarly situated whites but keeping the bias in the personal rating.[54] So even if one were to assume—erroneously, according to our analysis—that the personal rating was not biased, a substantial Asian American penalty remains. A more reasonable interpretation is that the reduction in the Asian American penalty implies that the personal rating accounts for a little less than half of the total Asian American penalty.

5.2 ALDC Applicants

Our preferred admissions model excludes ALDC applicants (recruited athletes, legacies, donor connections, children of faculty/staff) since there is ample evidence that the process works differently for them. Although Harvard claims the admissions process is the same for ALDC applicants, this doesn't appear to be the case. Athletes have a direct connection with an on-campus advocate (coaches), donor-connected applicants are literally on a special dean's interest list, and the dean of admissions directly reviews the files of legacies and children of faculty/staff.[55] Typical applicants receive none of these advantages. Thus, there is strong a priori evidence indicating that the process works differently for ALDC applicants.

In addition, the data reveal that other applicant characteristics influence admissions differently for ALDC applicants. In particular, the importance of academics and extracur-

[54]The Asian American penalty persists even when we include the overall rating as a control in a model with a slightly different sample and slightly different set of controls. See Table B.7.1 of Document 415-8. This estimate is not expressed as a marginal effect, so it is difficult to compare with other estimates. However, the Asian American coefficient is negative and significant.

[55]See Document 419-1, p. 41 and footnote 89 of Document 419-143 for documentation on the admissions process for athletes. Document 421-9 describes how donor-connected applicants are handled differently by the admissions office. For legacy applicants, the application reading procedures instruct that these files "should be read by [Admissions Dean Fitzsimmons] following the normal reading process if the decision might require special handling or if another reading might be helpful" (Trial Exhibit P001, p. 3). For children of faculty/staff, the reading procedures state that "All [faculty] and [staff] folders should be sent to [Dean Fitzsimmons] after the normal reading process has been completed." (Trial Exhibit P001, p. 3).

ricular activities is watered down. This is illustrated in Table 6 of Arcidiacono, Kinsler, and Ransom (forthcoming), where the coefficients on Harvard's academic and extracurricular ratings fall (especially at low ratings) when ALDCs are included in the admissions model. The differential treatment of these applicants can also be seen in Table F5. White legacy, donor-connected, and faculty/staff applicants (LDC) in the bottom 10% of the academic index distribution had a higher admit rate than the average white applicant. Further, in the bottom 10% of the academic index distribution, virtually all typical applicants are rejected. There were two African American typical admits in the bottom decile across the six admission cycles and no admits from any other racial group. This means that the admit rate for African American typical applicants in the bottom decile of the academic index is 0.03% as compared to an over 6% admit rate for white LDC applicants in the bottom decile.

In light of this evidence, there are two reasonable approaches for handling ALDC applicants. First, ALDC applicants can be added to the admissions model, but it is imperative to interact their special applicant status with many other applicant characteristics to allow for differential treatment. The second approach is to simply exclude ALDC applicants, and as we show in Appendix A, obtain consistent estimates of the Asian American penalty for typical applicants. We follow the second approach, where the only potential drawback is that we do not estimate the role that race plays in admissions for ALDC applicants. Since more than 97% of Asian American applicants are not ALDC, this seems like a reasonable tradeoff.

While we believe our decision to exclude ALDC applicants is appropriate, we have also estimated admissions models that include these special applicants. To be clear, these models do not interact special applicant status with academic and extracurricular characteristics as we advocate above, and as a result are likely to understate the penalty experienced by typical Asian American applicants.[56] When LDC applicants are added to our preferred admissions model, we find that the probability of admission for a typical Asian American applicant increases from 5.2% to 6.1% when treated as a white applicant (see Table 7.2R of Document 415-9).[57] Without LDC applicants in the model, the same thought experiment

[56]Here we are limited to models that are already part of the public record since we no longer have access to Harvard admissions data.

[57]When LDC applicants are added, we introduce indicator variables for legacy, double legacy, faculty or

yields an increase in admissions chances from 5.2% to 6.2%. While we have not estimated our preferred admissions model with recruited athletes, Tables B.7.1 and B.7.2 of Document 415-8 illustrate that in a slightly altered admissions model, the addition of ALDC applicants mildly reduces the negative impact of Asian American status for typical applicants.[58] This is similar to what we find when adding only LDC applicants to our preferred specification.

There are two important takeaways from the admissions models that include ALDC applicants. First, the estimated penalty for typical Asian American applicants remains large and statistically significant. Second, the estimates suggest that Asian American applicants who are ALDC are not penalized because of their race. These two results are not in conflict, as there is no ex-ante reason why discrimination should work the same across all groups of applicants. In fact, by belonging to one of the special applicant groups, Asian American applicants may be able to overcome stereotypes that hold typical Asian American applicants back.

The lack of a penalty against Asian American ALDC applicants should not diminish claims that Harvard employs admissions practices that discriminate against Asian Americans. More than 97% of Asian American applicants are not ALDC, meaning that nearly all Asian American applicants face an explicit penalty in admissions. Moreover, the very existence of ALDC preferences works to the detriment of the overwhelming majority of Asian Americans. ALDC applicants are predominantly white, and as we show in Arcidiacono, Kinsler, and Ransom (forthcoming), the elimination of either legacy or athlete preferences would increase the number of Asian American admits by more than 4%.

5.3 Scope for Omitted Variable Bias

Our estimate of the Asian American admissions penalty at Harvard is based on a logit model that includes more than 300 applicant characteristics. Included in these attributes are Harvard's internal ratings of applicants along a variety of dimensions, including extracurricular activities, recommendations from high school teachers and counselors, and alumni inter-

staff child, donor connection, interactions between legacy and race, and interactions between faculty/staff child or donor connection and race.

[58] See Section 8 of Document 415-9 for a detailed discussion of how this model differs from our preferred model.

views. Relative to previously published analyses of US admissions decisions, the richness of the available data is without rival.[59] However, the possibility remains that the estimated admissions penalty for Asian American applicants is not causal and instead reflects the impact of unobserved attributes that are more common among Asian American applicants relative to white applicants. We present evidence in the following sections that suggests this is unlikely.

5.3.1 Strength of Asian Americans on observables

While it is infeasible to directly test for omitted variable bias since the relevant attributes are by definition unobserved, we can examine the average strength of each racial group based on their "observed admissions index." If a group of applicants is strong on the observed attributes that predict admission, they are likely to be strong on unobserved attributes that predict admission. Using our estimated model, we construct an admissions index which assesses applicants' strengths based on how their observed characteristics translate into a probability of admission, after removing race and year effects. We then construct deciles of this admissions index, with higher deciles corresponding to stronger observed characteristics. The results for Asian American and white applicants are displayed in the first two columns of Table 8. Asian American applicants are stronger, with 13.1% of Asian American applicants in the top decile and only 10.5% of white applicants in the top decile; among the top two deciles, the Asian American share is 25.9%, while for whites it is 21.2%.[60]

The result that Asian American applicants are stronger on the observable characteristics associated with admissions is unsurprising given the incredible academic strength of this group.[61] However, Asian American applicants could be weaker along observed non-academic

[59] Various papers have explored the impact of race (Long, 2004; Arcidiacono, 2005; Antonovics and Backes, 2014) and legacy status (Espenshade, Chung, and Walling, 2004; Espenshade and Chung, 2005; Hurwitz, 2011) on US admissions decisions. Admissions models in these papers typically have access to only a handful of applicant attributes and typically have to estimate their average impact across multiple colleges. Bhattacharya, Kanaya, and Stevens (2017) focus on the impact of academic credentials on admissions decisions using detailed data from a selective UK university. This paper is closest to ours in the richness of the available data.

[60] While the Asian American strength is clear using any decile comparisons, focusing on the top two deciles is relevant because this is where most admitted students come from. For example, the results in Table 7 imply that over 93% of Asian American admissions come from those in the top 20% of the observed admissions index.

[61] Table 5.3 of Document 415-8 shows that over 51% of the class would be Asian American if admissions

dimensions. If this were the case, it would suggest that they might be weaker on unobserved non-academic dimensions. It is not clear whether the unobservables are disproportionately non-academic.[62] But we can assess whether Asian American applicants are weaker on the non-academic attributes related to admission. We construct a non-academic index by removing those characteristics that are explicitly academic in nature (e.g., test scores, grades, academic ratings) from the admissions index. Results are shown in the third and fourth columns of Table 8. In each of the top 4 deciles there is a larger share of Asian American applicants relative to white applicants. It is clear that, on non-academic measures, Asian American applicants are at least as strong as white applicants.

Included in the non-academic measures affecting admissions are certain attributes that are likely to favor Asian American applicants, such as disadvantaged and first-generation status. There are also non-academic attributes that likely harm the admissions chances of Asian American applicants, such as geography.[63] However, we can eliminate the impact of these attributes and construct a non-academic admissions index consisting only of Harvard's extracurricular, athletic, school support, and alumni ratings. The final two columns display the results when we construct the admissions index in this manner, and we still find that Asian American applicants are just as strong, if not stronger than white applicants.

5.3.2 Model Fit

Further limiting the scope for omitted variable bias is how well our preferred model fits the data. The Pseudo R^2—or McFadden's R^2—of our model is 0.56. While higher values of this measure indicate a better model fit, it does not have the same interpretation as the R^2 used in linear models except when either (i) the model explains the data completely (in which case they are both one) or (ii) the model only has an intercept term (in which case

were based entirely on the academic index.

[62] Despite the fact that our preferred model includes many academic measures, it is still likely that we fail to capture all dimensions of applicant academic success. For example, academic-related attributes, such as AP exams taken, AP scores, and academic awards such as winning the Harvard-MIT Mathematics Tournament (HMMT) are excluded from the model. We know that Asian Americans are stronger on AP exams, and are likely stronger on other unobserved academic measures.

[63] Document 415-9 section 9.3 shows that dockets (i.e. geographic areas) with a higher share of Asian Americans face a penalty. This can result from Harvard valuing geographic diversity in combination with Asian Americans being so competitive. A larger number of applicants are likely to come from areas that have higher shares of Asian Americans.

they are both zero). The classic citation on the relation between the two R^2's, McFadden (1979), suggests that 0.56 is well above what would be considered an excellent fit.[64] But the designation "excellent fit" is still not especially precise. To provide more evidence on the fit of the model, we first consider how our Pseudo R^2 translates into an R^2 of the latent index and then examine the accuracy of the admissions model.

We can link the two R^2 measures by returning to the underlying model of the admissions process. Namely, when an applicant's latent index, Y_i^*, exceeds some threshold τ, they are admitted. Denote the observed part of this index as $AI_i = X_i\beta$, implying:

$$Y_i^* = AI_i + \epsilon_i. \tag{2}$$

Following McKelvey and Zavoina (1975) and expanded upon by Veall and Zimmermann (1996), we can calculate how much of the observables—as measured by AI_i—explain the (implicit) scoring of Harvard's applicants Y_i^*. To calculate the implied R^2 associated with (2), we simulate AI_i's and ϵ_i's that are consistent with (i) an overall admit rate of 5.45% and (ii) a Pseudo R^2 of 0.56. The simulation of the ϵ's entails draws from a logistic distribution as that is what was used to generate the model estimates.

What is unknown is the distribution of AI_i.[65] Assuming that the distribution of AI_i follows a normal distribution and matching the overall admit rate and the Pseudo R^2 of the model results in an implied R^2 of 0.8 for the latent index. However, the implied R^2 will be sensitive to the distribution chosen for AI_i. In Appendix D, we show that there is additional information in the public reports that is helpful in recovering the distribution of AI_i. Incorporating this additional information suggests that an implied R^2 of 0.8 is conservative, though also relies on assumptions about the the tails of the distribution of AI_i.

We use a similar approach to calculate the accuracy of the admissions model which, as we

[64] McFadden (1979), p. 307, states that

> Those unfamiliar with the ρ^2 index should be forewarned that its values tend to be considerably lower than those of the R^2 index and should not be judged by the standards for a "good fit" in ordinary regression analysis. For example, values of 0.2 to 0.4 for ρ^2 represent an excellent fit.

The ρ^2 referred to here later became known as McFadden's R^2, or the Pseudo R^2.

[65] This distribution could be calculated using the underlying data. However, the information needed to do so is not in the public record.

show in Appendix D, is less sensitive to the choice of the distribution of AI_i. Namely, given the distributions of AI_i and ϵ_i, we simulate admissions decisions and ask how well the model predicts admissions decisions based on AI_i alone. The accuracy for admits is then what share of the 5.45% of simulated admits based on AI_i and ϵ_i are in the top 5.45% of the AI_i distribution. The accuracy for admits is over 64%, remarkably high given that only 5.45% of applicants are admitted. We can similarly compute the accuracy of rejects, i.e. what share of the simulated rejects are in the bottom 94.55% of the AI_i distribution. Predicting rejection is much easier since so few applicants are admitted, and our corresponding accuracy of rejects is over 99%.

The exceptional fit of the model leaves only a limited amount of unobserved information. Given that Asian Americans are stronger overall on variables that account for a substantial amount of Harvard's admissions decisions, it would be remarkable if they were significantly worse on the small portion of characteristics that are unobservable.

6 Conclusion

The perception that Asian Americans are discriminated against in elite college admissions has led college consultants to "make them less Asian when they apply" (English, 2015). Using data made public from the *SFFA v. Harvard* case, we show that this perception is justified for almost all Asian American applicants.[66]

The discrimination manifests itself both in a direct penalty in admissions, but also in an Asian American penalty in some of Harvard's ratings. Asian Americans are stronger than white applicants on the observables associated with each of the ratings with the exception of the athletic rating, which was not modeled. Yet, on ratings like the personal and overall rating, Asian Americans receive lower ratings.[67] For example, Asian Americans would see 20% higher odds of receiving a 2 or better on the personal rating if they were treated as white applicants. These odds would almost double if they were treated as African American applicants.

[66] The exception is the less than 3% of Asian American applicants who are ALDC.

[67] At the same time, underrepresented minorities receive a bump on these ratings despite being substantially worse on the observables associated with each of the ratings.

These penalties against Asian Americans in the ratings also translate to penalties in admissions. Using whites as a base, our preferred model shows an average marginal effect -1 percentage point for being Asian American. This implies a 19% penalty given the admission rate for typical Asian Americans was slightly over 5% for the period we analyze. This admissions penalty is likely an understatement for the following reasons: (i) Asian Americans are stronger than whites on the observables associated with admission; and (ii) there is evidence of bias against Asian Americans in some of the other ratings that are included in the model.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1):151–184. 5, A18

Antonovics, Kate and Ben Backes. 2014. "The Effect of Banning Affirmative Action on College Admissions Policies and Student Quality." *Journal of Human Resources* 49 (2):295–322. 25

Anwar, Shamena and Hanming Fang. 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 96 (1):127–151. 4

Arcidiacono, Peter. 2005. "Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica* 73 (5):1477–1524. 16, 25

Arcidiacono, Peter, Josh Kinsler, and Tyler Ransom. forthcoming. "Legacy and Athlete Preferences at Harvard." *Journal of Labor Economics* . 2, 8, 12, 15, 23, 24, A16

Arnold, David, Will Dobbie, and Crystal S Yang. 2018. "Racial Bias in Bail Decisions." *Quarterly Journal of Economics* 133 (4):1885–1932. 4

Bhattacharya, Debopam, Shin Kanaya, and Margaret Stevens. 2017. "Are University Admissions Academically Fair?" *Review of Economics and Statistics* 99 (3):449–464. 4, 25

Blau, Francine D. and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3):789–865. 4

Bordon, Paola and Chao Fu. 2015. "College-Major Choice to College-Then-Major Choice." *Review of Economic Studies* 82 (4):1247–1288. 4

Canay, Ivan A, Magne Mogstad, and Jack Mountjoy. 2020. "On the Use of Outcome Tests for Detecting Bias in Decision Making." Working Paper 27802, National Bureau of Economic Research. 4

Day 14 Trial Transcript. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialTranscripts/20181101-Day-14.pdf. 5, A31

Day 4 Trial Transcript. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialTranscripts/20181018-Day-4.pdf. 3, A31

Ding, Weili and Steven F Lehrer. 2007. "Do Peers Affect Student Achievement in China's Secondary Schools?" *Review of Economics and Statistics* 89 (2):300–312. 4

Document 1. 2020. In *USA v. Yale University*. Civil Action No. 3:20-cv-01534 (D. Conn). URL https://www.justice.gov/opa/press-release/file/1326306/download. Complaint. 3, 10

Document 160-1. 2019. In *Students for Fair Admissions, Inc. v. University of North Carolina*. Civil Action No. 1:14-cv-00954-LCB-JLW (M.D.N.C.). URL https://samv91khoyt2i553a2t1s05i-wpengine.netdna-ssl.com/wp-content/uploads/2019/02/160-1.pdf. Plaintiff's Expert Witness Report. 21

Document 415-8. 2017. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://docs.justia.com/cases/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/415/1.html. Plaintiff Expert Witness Opening Report. 7, 9, 22, 24, 25, A16, A31

Document 415-9. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://docs.justia.com/cases/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/415/2.html. Plaintiff Expert Witness Rebuttal Report. 6, 12, 16, 19, 21, 23, 24, 26, 35, 36, 37, 38, 39, 40, 41, A15, A16, A17, A18, A31, A35

Document 419-1. 2015. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://www.courtlistener.com/recap/gov.uscourts.mad.165519/gov.uscourts.mad.165519.419.1_2.pdf. Deposition of Marlyn McGrath. 22, A31

Document 419-141. 2017. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://projects.iq.harvard.edu/files/diverse-education/files/expert_report_-_2017-12-15_dr._david_card_expert_report_updated_confid_desigs_redacted.pdf. Defendant Expert Witness Opening Report. 5, A31

Document 419-143. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://projects.iq.harvard.edu/files/diverse-education/files/expert_report_rebuttal_as_filed_d._mass._14-cv-14176_dckt_000419_037_filed_2018-06-15.pdf. Defendant Expert Witness Rebuttal Report. 22, A31

Document 421-9. 2017. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://www.courtlistener.com/recap/gov.uscourts.mad.165519/gov.uscourts.mad.165519.421.9_7_1.pdf. Deposition of William Fitzsimmons. 13, 22, A31

English, Bella. 2015. "To Get into Elite Colleges, Some Advised to 'Appear Less Asian'." *Boston Globe* URL https://www.bostonglobe.com/lifestyle/2015/06/01/college-counselors-advise-some-asian-students-appear-less-asian/Ew7g4JiQMiqYNQlIwqEIuO/story.html. Accessed February 14, 2020. 9, 28

Espenshade, Thomas J. and Chang Y. Chung. 2005. "The Opportunity Cost of Admission Preferences at Elite Universities." *Social Science Quarterly* 86 (2):293–305. 4, 25

Espenshade, Thomas J., Chang Y. Chung, and Joan L. Walling. 2004. "Admission Preferences for Minority Students, Athletes, and Legacies at Elite Universities." *Social Science Quarterly* 85 (5):1422–1446. 4, 5, 25

Fuchs, Chris. 2019. "After Harvard affirmative action decision, Asian American students rethink college applications." *NBC News* URL https://www.nbcnews.com/news/asian-america/after-harvard-affirmative-action-decision-asian-american-students-rethink-college-n10 Accessed February 28, 2020. 2

Golden, Daniel. 2006. *The Price of Admission: How America's Ruling Class Buys Its Way into Elite Colleges – and Who Gets Left Outside the Gates*. New York: Crown Publishers. 2

Goodman, Joshua, Oded Gurantz, and Jonathan Smith. 2020. "Take Two! SAT Retaking and College Enrollment Gaps." *American Economic Journal: Economic Policy* 12 (2):115–158. URL https://www.aeaweb.org/articles?id=10.1257/pol.20170503. 10

Hurwitz, Michael. 2011. "The Impact of Legacy Status on Undergraduate Admissions at Elite Colleges and Universities." *Economics of Education Review* 30 (3):480–492. 4, 25

Jia, Ruixue and Hongbin Li. 2017. "Access to Elite Education, Wage Premium, and Social Mobility: The Truth and Illusion of China's College Entrance Exam." Working paper, Stanford University. 4

Ketel, Nadine, Edwin Leuven, Hessel Oosterbeek, and Bas van der Klaauw. 2016. "The Returns to Medical School: Evidence from Admission Lotteries." *American Economic Journal: Applied Economics* 8 (2):225–254. 4

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *Quarterly Journal of Economics* 131 (3):1057–1111. 4

Knowles, John, Nicola Persico, and Petra Todd. 2001. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109 (1):203–229. 4

Krauth, Brian. 2016. "Bounding a Linear Causal Effect Using Relative Correlation Restrictions." *Journal of Econometric Methods* 5 (1):117–141. 5

Long, Mark C. 2004. "Race and College Admissions: An Alternative to Affirmative Action?" *Review of Economics and Statistics* 86 (4):1020–1033. 4, 25

McFadden, Daniel. 1979. "Quantitative Methods for Analysing Travel Behaviour: Some Recent Developments." In *Behavioural Travel Modeling*, edited by David A. Hensher and Peter R. Stopher, chap. 13. Croom Helm Ltd. 5, 27

McKelvey, Richard D. and William Zavoina. 1975. "A Statistical Model for the Analysis of Ordinal Level Dependent Variables." *The Journal of Mathematical Sociology* 4 (1):103–120. URL https://doi.org/10.1080/0022250X.1975.9989847. 27

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37 (2):187–204. 5

Trial Exhibit DX 024. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/D024.pdf. 2012 Casebook Discussion Guide. A31

Trial Exhibit DX 042. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/D042.pdf. Demographic Breakdown of Applicants, Admits, and Matriculants. 6, 7, A31

Trial Exhibit DX 680. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/D680.pdf. Primary extracurricular activity by race. A31

Trial Exhibit P001. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P001.pdf. Harvard Reading Procedures for Class of 2018. 7, 8, 18, 22, A19, A31

Trial Exhibit P009. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P009.pdf. Office of Institutional Research report, "Admissions Part II: Subtitle". 4, A31

Trial Exhibit P028. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P028.pdf. Office of Institutional Research report, "Demographics of Harvard College Applicants". 4, A31

Trial Exhibit P164. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P164.pdf. One-pager. 7, A31

Trial Exhibit P555. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P555.pdf. Office of Civil Rights Report. 12, A31

Trial Exhibit P621. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P621.pdf. Ratings frequencies for baseline sample. 11, 36, A31

Trial Exhibit P633. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P633.pdf. Harvard Reading Procedures for Class of 2023. 7, A19, A31

US Census Bureau. 2020. "Historical Income Tables: Households [Data file]." URL https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html. Accessed Jan 27 2021. 9

Veall, Michael R. and Klaus F. Zimmermann. 1996. "Pseudo-R2 Measures for Some Common Limited Dependent Variable Models." *Journal of Economic Surveys* 10 (3):241–259. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6419.1996.tb00013.x. 27

Wong, Alia. 2019. "4 Myths Fueling the Fight Over NYCs Exclusive High Schools." *The Atlantic* URL https://www.theatlantic.com/education/archive/2019/03/stuyvesant-admissions-controversy-fact-or-fiction/585460/. Accessed Jan 27, 2021. 9

Zimmerman, Seth D. 2014. "The Returns to College Admission for Academically Marginal Students." *Journal of Labor Economics* 32 (4):711–754. 4

———. 2019. "Elite Colleges and Upward Mobility to Top Jobs and Top Incomes." *American Economic Review* 109 (1):1–147. 4

Figures and Tables

Table 1: Summary Statistics of White and Asian American Applicants and Admits

	White			Asian American		
	Reject	Admit	Total	Reject	Admit	Total
Panel A: Demographics						
Admitted	0.00	100.00	4.89	0.00	100.00	5.13
Female	45.75	43.14	45.62	49.12	52.65	49.30
Disadvantaged	5.94	14.61	6.36	10.26	21.86	10.85
First-generation college	4.29	4.05	4.28	7.98	9.65	8.07
Applied for fee waiver	8.00	12.15	8.20	12.88	18.39	13.16
Applied for financial aid	73.83	72.17	73.75	76.37	77.27	76.41
Mother's education: MA or higher	37.86	46.24	38.27	37.63	44.78	38.00
Father's education: MA or higher	46.36	52.38	46.65	54.89	59.60	55.13
Panel B: Academic Preparation						
SAT1 math (z-score)	0.12	0.56	0.15	0.41	0.77	0.43
	(0.82)	(0.50)	(0.81)	(0.73)	(0.37)	(0.72)
SAT1 verbal (z-score)	0.31	0.72	0.33	0.31	0.74	0.33
	(0.76)	(0.43)	(0.75)	(0.80)	(0.41)	(0.79)
SAT2 avg (z-score)	-0.01	0.58	0.03	0.32	0.81	0.35
	(0.86)	(0.50)	(0.85)	(0.82)	(0.38)	(0.81)
Standardized high school GPA (z-score)	0.17	0.50	0.18	0.21	0.52	0.22
	(0.86)	(0.52)	(0.85)	(0.82)	(0.47)	(0.81)
Academic index (z-score)	0.16	0.76	0.19	0.39	0.91	0.42
	(0.80)	(0.38)	(0.79)	(0.78)	(0.32)	(0.77)
Academic index percentile	0.52	0.75	0.53	(0.78)	0.83	0.63
	(0.26)	(0.19)	(0.26)	(0.27)	(0.16)	(0.27)
Number of AP tests taken	4.08	5.91	4.16	5.60	7.50	5.68
	(3.91)	(3.85)	(3.93)	(4.07)	(3.38)	(4.06)
Average score of AP tests	4.39	4.74	4.41	4.48	4.82	4.50
	(0.59)	(0.34)	(0.58)	(0.56)	(0.28)	(0.55)
N	54,768	2,814	57,582	38,343	2,072	40,415

Source: Table B.3.1R of Document 415-9.

Notes: Data restricted to typical (non-ALDC) applicants from the Classes of 2014–2019. Standard deviations in parentheses. AP scores are only available for a subset of the years.

Table 2: Shares and Admission Rates of Applicants by Academic Index Decile and Race

	Number of Applicants		Share of Applicants		Admit Rate	
Decile	White	Asian American	White	Asian American	White	Asian American
1	2,822	1,511	4.91	3.75	0.00	0.00
2	4,404	2,045	7.67	5.07	0.39	0.20
3	6,073	2,644	10.57	6.56	0.56	0.64
4	6,359	3,020	11.07	7.49	1.82	0.86
5	7,658	3,874	13.33	9.61	2.57	1.86
6	5,924	3,614	10.31	8.97	4.20	2.49
7	7,053	4,527	12.28	11.23	4.79	3.98
8	6,478	5,316	11.28	13.19	7.53	5.12
9	5,717	6,532	9.95	16.21	10.77	7.55
10	4,963	7,225	8.64	17.92	15.27	12.69

Source: Authors' calculations from data presented in Table 5.1R of Document 415-9.

Notes: Share columns sum to 100 within each group. Data restricted to typical (non-ALDC) applicants from the Classes of 2014–2019.

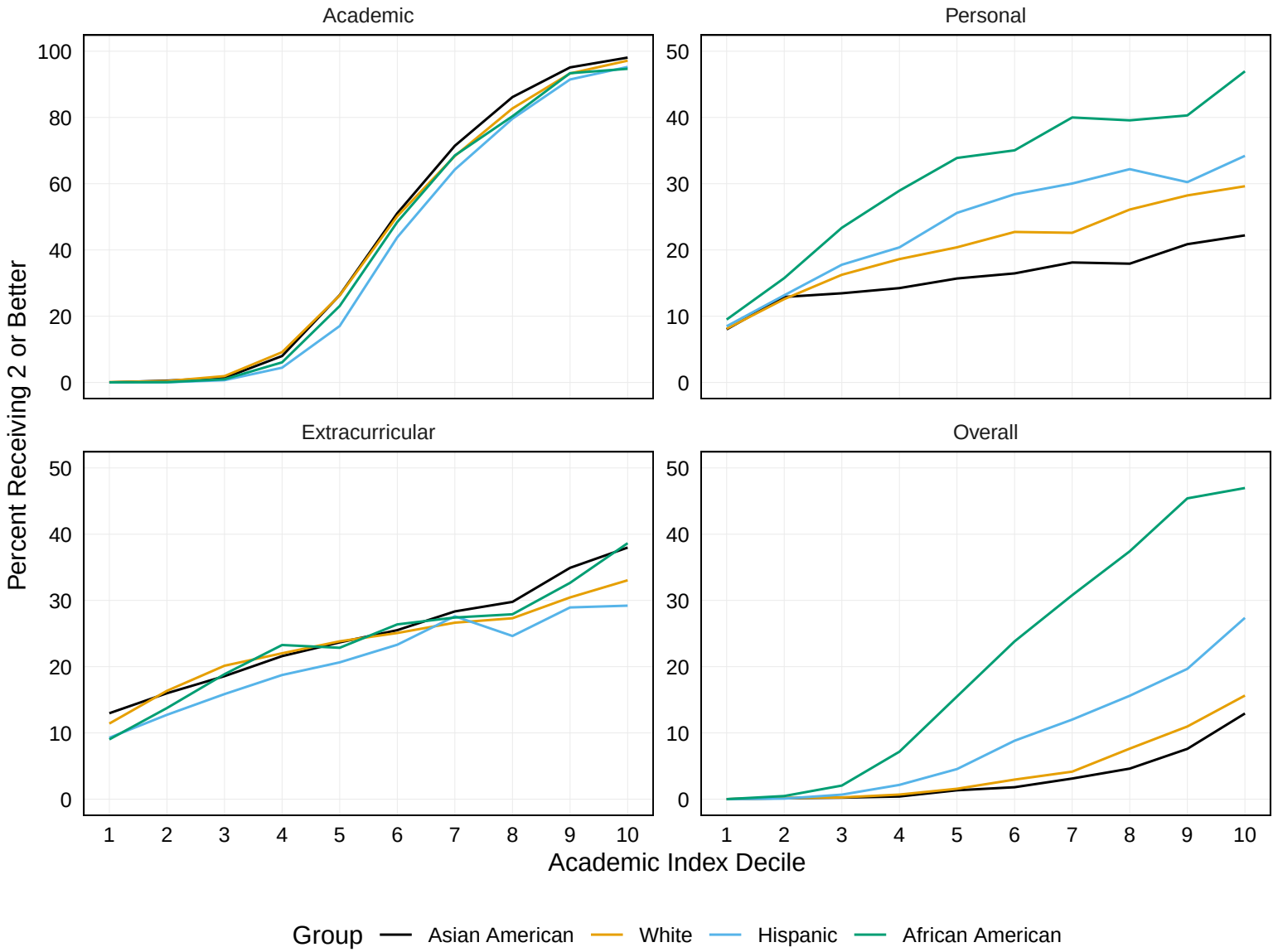
Table 3: Share of Applicants Receiving a 2 or Better on Application Ratings

Rating	White	Asian American
Overall	4.43	4.84
Academic	45.29	60.21
Extracurricular	24.35	28.23
Athletic	12.79	4.81
Personal	21.27	17.64
Teacher 1	30.42	30.79
Teacher 2	27.13	27.41
Counselor	25.28	25.12
Alumni Personal	49.92	50.33
Alumni Overall	36.49	40.89

Source: Authors' calculations from data presented in Trial Exhibit P621.

Notes: Those with missing ratings are coded as not having received a 2 or better. Data restricted to typical (non-ALDC) applicants from the Classes of 2014–2019.

Figure 1: Percent Receiving 2 or Better on Various Ratings by Race and Academic Index Decile



Source: Authors' calculations from Tables 5.4R, 5.6R and 5.7R of Document 415-9.

Table 4: Asian American Ratings and Admit Rate in Top Decile of Academic Index, Compared to Other Race Groups in Top Decile

		Comparison with Whites			Comparison with African Americans			Comparison with Hispanics		
Rating/Outcome	Asian American Rate in Top AI Decile	Difference in Top AI Decile	Pct Increase/ Decrease in Top AI Decile	Lowest Decile with Higher Rate	Difference in Top AI Decile	Pct Increase/ Decrease in Top AI Decile	Lowest Decile with Higher Rate	Difference in Top AI Decile	Pct Increase/ Decrease in Top AI Decile	Lowest Decile with Higher Rate
Personal	22.20	7.42	33.42	6	24.77	111.57	3	12.01	54.10	5
Counselor	38.34	6.29	16.41	9	10.90	28.44	9	6.66	17.37	10
Teacher 2	41.90	5.21	12.44	10	8.86	21.15	9	7.84	18.71	10
Teacher 1	46.64	3.53	7.56	10	8.66	18.57	9	2.83	6.07	10
Alumni Personal	63.61	1.37	2.15	10	9.87	15.52	7	7.44	11.70	10
Alumni Overall	63.10	0.03	0.04	10	3.57	5.65	10	1.37	2.18	10
Academic	98.08	-0.92	-0.94	-10	-3.38	-3.45	-10	-2.81	-2.87	-10
Extracurricular	37.98	-4.93	-12.99	-9	0.66	1.73	10	-8.77	-23.09	-8
Admit	12.69	2.58	20.34	10	43.37	341.70	4	18.62	146.74	6
Overall	12.93	2.71	20.95	10	34.04	263.34	5	14.44	111.71	8

Source: Authors' calculations from Tables 5.2R and 5.4–5.7R of Document 415-9.

Notes: All results are conditional on being in the highest academic index decile. Negative signs in front of deciles (in columns 4, 7 and 10) emphasize that Asian Americans perform better than the given group on that particular rating. Note that this is only ever true for the academic and extracurricular ratings.

Table 5: Selected Coefficients, Admissions Models

	(1)	(2)	(3)	(4)	(5)	(6)
African American	0.531	2.417	2.671	2.851	3.772	3.876
	(0.040)	(0.050)	(0.074)	(0.078)	(0.105)	(0.112)
Hispanic	0.425	1.273	1.286	1.339	1.959	2.027
	(0.039)	(0.044)	(0.063)	(0.067)	(0.085)	(0.091)
Asian American	0.057	-0.434	-0.565	-0.378	-0.466	-0.330
	(0.032)	(0.035)	(0.052)	(0.055)	(0.070)	(0.074)
Female	-0.044	0.254	0.228	0.271	0.163	0.141
	(0.025)	(0.027)	(0.064)	(0.088)	(0.110)	(0.116)
Disadvantaged	1.183	1.257	1.497	1.606	1.660	1.535
	(0.042)	(0.048)	(0.071)	(0.108)	(0.138)	(0.147)
Early Action	1.616	1.456	1.371	1.348	1.410	1.440
	(0.032)	(0.035)	(0.055)	(0.084)	(0.104)	(0.110)
Academic Rating=4					-3.990	-3.915
					(0.626)	(0.633)
Academic Rating=2					1.425	1.941
					(0.090)	(0.128)
Academic Rating=1					4.094	5.122
					(0.156)	(0.185)
Extracurricular Rating=4					-1.301	-1.122
					(0.393)	(0.408)
Extracurricular Rating=2					1.990	1.810
					(0.082)	(0.108)
Extracurricular Rating=1					4.232	4.215
					(0.169)	(0.187)
Athletic Rating=4					-0.182	-0.043
					(0.038)	(0.041)
Athletic Rating=2					1.368	1.354
					(0.114)	(0.155)
N	142,728	142,700	142,700	136,061	128,422	128,082
Pseudo R Sq.	0.078	0.260	0.262	0.283	0.556	0.604
Demographics	Y	Y	Y	Y	Y	Y
Academics	N	Y	Y	Y	Y	Y
Race and Gender Interactions	N	N	Y	Y	Y	Y
HS and NBHD Variables	N	N	N	Y	Y	Y
Ratings (excluding Personal)	N	N	N	N	Y	Y
Personal Rating	N	N	N	N	N	Y

Source: Data presented in Table B.7.1R of Document 415-9.

Notes: All models include year indicators and year interactions. Standard errors reported below each coefficient in parentheses. In models (3)-(6), the race coefficients reflect preferences for male, non-disadvantaged students. The excluded ratings categories are a 3. A full list of controls is available in Appendix B.

Table 6: Probability of Admission (%) for an Asian American if Treated Like a White Applicant

	Baseline Probability (%)			
Group	1.00	5.00	10.00	25.00
Asian, male, not disadvantaged	1.58	7.74	15.04	34.69
Asian, female, not disadvantaged	1.26	6.25	12.34	29.70
Asian, male, disadvantaged	1.37	6.77	13.29	31.51
Asian, female, disadvantaged	1.10	5.46	10.87	26.78

Source: Calculations based on coefficients listed in Table 5 and formula given in Equation (1).

Table 7: The Asian American Penalty at Different Admissions Deciles

Admissions Index Decile	Marginal Effect	Admission Prob. w/ Penalty	Admission Prob. no Penalty	Pct. Increase if Penalty Removed
5 and Below	-0.02%	0.04%	0.06%	40.24%
6	-0.13%	0.32%	0.44%	39.81%
7	-0.31%	0.77%	1.08%	39.98%
8	-0.78%	2.03%	2.82%	38.63%
9	-2.45%	7.01%	9.46%	34.98%
10	-6.19%	41.68%	47.87%	14.84%

Source: Table 9.1 of Document 415-9.

Notes: Admissions index decile refers to the ranking of Asian-American applicants by their estimated admission index (i.e. the controls times their coefficients), absent admissions cycle.

Table 8: Distribution of White and Asian American Applicants (%) by Strength on Observed Factors Affecting Admission

	Admissions Index		Non-Academic Admissions Index		Non-Academic Ratings Admissions Index	
Decile	Asian American	White	Asian American	White	Asian American	White
5 or lower	38.1	45.8	46.6	48.0	43.9	45.7
6	11.3	11.1	10.4	10.6	11.3	10.5
7	12.0	11.2	10.7	10.4	10.6	10.9
8	12.8	10.7	10.9	10.4	11.1	10.8
9	12.8	10.7	11.0	10.3	11.7	10.8
10	13.1	10.5	10.4	10.3	11.3	11.3

Source: Tables 7.3R, 7.4R, and 7.5R of Document 415-9.

Notes: Numbers indicate the percentage of applicants within each cell. Each column sums to 100.

Decile refers to the ranking of typical applicants on the given dimension of their estimated admissions index. The admissions index includes all covariates in the admissions model except race and the admissions cycle. The non-academic admissions index excludes test scores, grades and academic ratings from the admissions index. The non-academic ratings admissions index excludes all admissions model covariates except the following Harvard ratings: extracurricular, athletic, school support, and alumni ratings.

Online Appendices

A Modeling Admissions and the Relevant Sample

We focus on estimating the Asian American penalty among typical applicants.[A1] This decision reflects a desire to make comparisons of similarly-situated applicants. In this section we expand on this idea, showing that consistent estimates of a penalty against typical Asian American applicants can be recovered from the subset of applications that are typical (i.e. not ALDC). Further, incorporating ALDC applicants into the estimation in a meaningful way requires stronger assumptions in order to recover consistent estimates of a penalty against typical Asian American applicants.[A2] As we show, these assumptions are violated in the Harvard data.

In order to determine whether Harvard is discriminating against *typical* Asian Americans, it is only necessary to consider typical applicants. Consider a model of admissions where applicants compete for a fixed number of slots, N. Consider a latent index Y_i^* that represents Harvard's perception of the quality of applicant i. All applicants above some threshold τ are admitted: if $Y_i^* > \tau$ then $Y_i = 1$ and the applicant is admitted, otherwise the applicant is rejected. That all applicants face the same admissions threshold is without loss of generality: any preferences (e.g. for ALDC applicants or particular racial groups) can be folded into Y_i^*.[A3] The threshold τ is set to ensure the number of admits equals N.

The competition for slots manifests itself through τ. So even though, in principle, all students are competing against one another for a limited number of slots, it is still possible to estimate models of admission on subsets of the applicants, as any competitive effects will be reflected in τ.

Consider the subset of applicants that are typical (i.e. not ALDC). Further decompose

[A1] Discrimination may occur against subgroups of applicants. For example, married women or women with children may be treated differently in the labor market than single women. Given ALDC applicants have a clear tie to Harvard, it would make sense that discrimination would be more likely among non-ALDC applicants.

[A2] Incorporating ALDC applicants by interacting ALDC with every variable effectively allows the admissions process to operate differently for typical applicants and ALDC applicants. The addition of ALDC applicants in the fully interacted model then would have no effect on the estimates of racial discrimination against typical applicants.

[A3] Preferences for balancing factors such as initial major interest or geography can also be included in the latent index. For example, if there are few humanities applicants in a particular admissions cycle, then humanities applicants may see higher latent indexes all else equal.

the latent index Y_i^* for these applicants into the sum of three parts:

$$Y_i^* = \alpha A_i + X_i\beta + \epsilon_i \tag{A.1}$$

(*i*) a part due to being Asian American, $A_i = 1$; (*ii*) a part due to other observables, X_i; and (*iii*) a part due to unobservables, ϵ_i.

To fix ideas, consider the case where ϵ_i follows a logistic distribution and is uncorrelated with A_i and X_i. In this case, a logit model yields consistent estimates of the parameters α and β up to a scale parameter, where the scale parameter embeds the variance of ϵ.[A4] Embedded in β will be a constant term that can be interpreted as a scaled version of τ.

Now suppose we add ALDC applicants to the data set, adding controls for their ALDC status to X_i to reflect any preferences these applicants receive. Suppose the same conditions hold as before: ϵ_i follows a logistic distribution and is uncorrelated with A_i and X_i. In this case, the additional observations increase the statistical power of the model. They do not, however, affect consistency: consistent estimates of the parameters can be obtained from a subset of the applicants.

However, if the Asian American coefficient substantially changes when ALDC applicants are added and this change is statistically significant, this suggests the model is misspecified. Either the effect of being Asian American differs for ALDC applicants, or other characteristics operate differently for ALDC applicants that in turn affect the coefficient on Asian American. Note that this is *not* a case of adding controls and having the coefficient change (i.e. omitted variable bias), but of adding observations.

One potential fix would be to allow the effect of being Asian American to vary between ALDC and typical applicants. Denoting $S_i = 1$ ($S_i = 0$) if the applicant was (was not) ALDC, we can specify the index as:

$$Y_i^* = \sum_{s=0}^{1} I(S_i = s)\alpha_s A_i + X_i\beta + \epsilon_i \tag{A.2}$$

But again, if the estimate of α_0 is substantially different when ALDCs are included, this

[A4] Note that the marginal effect of A_i is not affected by the normalization on the scale parameter.

suggests misspecification: the other variables matter in a different way for ALDC applicants, which in turn affects the Asian American coefficient.

A more general model—where ALDC status is interacted with all variables—would be:

$$Y_i^* = \sum_{s=0}^{1} I(S_i = s)\left(\alpha_s A_i + X_i \beta_s\right) + \epsilon_i \tag{A.3}$$

Note that this model implicitly builds in differences in the unobservables as well. By estimating separate coefficients for ALDC applicants, it allows for the possibility that the variance of ϵ is different for ALDC applicants. The coefficients for each group are all estimated relative to the underlying variances of their unobservables.

The model given in (A.3) can be estimated as two separate logits. These two logits will yield identical estimates to one where all the coefficients are estimated at once. Hence, for the purpose of estimating α_0, estimating a logit only on typical applicants is sufficient.

As discussed in section 5.2, the effects of race on admissions *and* the effects of X_i on admissions differ for ALDC applicants. Hence our primary focus is on models that exclude ALDC applicants in order to correctly measure discrimination against typical Asian American applicants. However, the inclusion of ALDC applicants still leads to a significant penalty against typical applicants, attenuating the average marginal effect by 0.1 percentage points.

B Admissions Controls

The list below describes the full set of variables we include in each of our admissions models. Our preferred specification is Model 5. This list comes from Figure 7.1 of Document 415-8, with additional information reported in Section 8.1 of Document 415-9.

- Model 1: Race/ethnicity, female, disadvantaged, application waiver, applied for financial aid, first generation college student, mother's education indicators, father's education indicators, year effects, docket-by-year effects, early action, intended major
- Model 2: Model 1 plus SAT math,* SAT verbal,* SAT2 average,* missing SAT2 average times race/ethnicity, converted GPA,* academic index,* academic index squared times academic index greater than zero, academic index squared times academic index less than zero, flag for converted GPA=35 (* indicates variable was z-scored)
- Model 3: Model 2 plus female times intended major, female times race/ethnicity, race/ethnicity times disadvantaged, race times early action
- Model 4: Model 3 plus College Board variables on the characteristics of applicant high schools and home neighborhoods (many are interacted with an indicator for whether the state is an SAT majority state), whether the mother or father is deceased, whether a parent attended an Ivy League university (other than Harvard), whether a parent attended graduate school at Harvard, the type of high school the applicant attended, an indicator for rural, an indicator for being a permanent resident, and year interacted with indicators for disadvantaged, first-generation, early action, financial aid, permanent resident, intended major, flag for converted GPA=35, and missing SAT2 average
- Model 5 (Preferred): Model 4 plus indicators for each category of the academic, extracurricular, athletic, teacher 1, teacher 2, counselor, alumni personal, and alumni overall ratings, interactions with missing alumni overall rating and race/ethnicity, indicators for whether the applicant had each possible combination of a two or better on Harvard's academic, extracurricular, and athletic profile ratings, indicators for whether

the applicant had two or three 2's or better on their school support measures, and an indicator for whether the applicant had 2's or better on both of the alumni ratings

- Model 6: Model 5 plus indicators for each category of the personal rating, and indicators for whether the applicant had a two or better on Harvard's personal rating in combination with a two or better on the academic, extracurricular, and athletic profile ratings

C Discrimination in Harvard Ratings

A key advantage of working with the Harvard data to ascertain racial preferences or penalties in admissions is the availability of Harvard's own internal ratings for each applicant. These ratings can potentially capture important applicant attributes that are unobserved to the researcher but correlated with applicant race. Including these ratings in a model of admissions will thus reduce the scope for omitted variable bias. However, if Harvard's applicant ratings also encompass racial preferences, then they would be inappropriate controls in an admissions model aimed at estimating the role of applicant race.

In this section, we more thoroughly investigate whether Harvard's ratings incorporate racial preferences, making them improper controls in an admissions model. To isolate the effect of race in applicant ratings, we estimate a series of ordered logit regressions where the outcome is a rating of interest, say the extracurricular rating, and the key controls include applicant race, gender, test scores, disadvantaged status, intended major, geography, neighborhood characteristics, and high school characteristics.[A5] Importantly, in each of the ratings models, we condition on all of the other Harvard ratings, excluding the personal and overall ratings. We exclude the personal and overall ratings since, as shown in Figure 1 and additional evidence below will show, they directly incorporate racial preferences. But if there is bias against Asian Americans in some of the other ratings, controlling for these ratings will lead to an under-estimate of any bias against Asian Americans. Hence our test is quite stringent.[A6]

Selected coefficients from the ordered logit ratings models are presented in Appendix Tables F6 and F7. For each rating, we present two models, one that is fairly sparse in terms of controls and our full model that contains a broad array of applicant attributes, including other ratings.[A7] We present both models to illustrate how the race coefficients change as controls are added—information which can be used as a guide to how unobservables correlate with race. Broadly speaking, the racial categories for African American, Hispanic, and

[A5] There are no estimates of the athletic rating in the public record but we do control for it in our preferred ratings models.

[A6] For a full description of the variables included in the ratings models, see Document 415-8 and Document 415-9.

[A7] Models with subsets of the controls can be found in Document 415-9 Tables B.6.1R–B.6.2R.

Asian American are statistically significant across most of the ratings models. However, the magnitudes, signs, and patterns in the coefficients as controls are added are quite different across Harvard's ratings.

In our full model, the Asian American coefficient is positive and significant for three of the ratings: the academic, extracurricular, and the alumni overall rating. But in each of these cases the Asian American coefficient—and indeed all the race coefficients—is much smaller in the specification of the full model than in the sparse model.[A8] A very different pattern emerges for the personal and overall rating. In our preferred models of the personal and overall rating, the Asian American coefficient is large and negative while the African American coefficient is large and positive, with the gaps growing between the sparse and full models. To put the magnitude of these coefficients in context, we calculate the change in the probability of obtaining a 2 or better for an Asian American if we were to switch their race, all else equal. If Asian American applicants were treated as African American (white) applicants, their probability of obtaining a 2 or better on the personal rating would increase by 59% (21%).[A9]

The large race coefficients in the estimated models for the personal and overall ratings suggest that racial preferences are an important factor in the assignment of these ratings. Further evidence that these coefficients arise from racial preferences can be seen in how the coefficients change between our sparse and full models. Appendix Figure F1 illustrates how the estimated racial gap between Asian American and African American applicants changes as controls are added for the academic, extracurricular, personal, and overall ratings. For the academic and extracurricular models, when we add controls the racial gap between these two groups shrinks towards zero. Note that, in our preferred ratings model, we interact race with gender and disadvantaged status. As a result, we show the male/female range for the estimated gap by disadvantaged status.

A very different pattern emerges for the personal and overall ratings. For the personal rating, the estimated racial gap between Asian American and African American applicants expands significantly as more controls are added. This is especially true for non-disadvantaged

[A8] Although here we are referring to the base coefficient, the finding of a diminished effect of race holds for any set of interactions as well.

[A9] See Table 6.1R in Document 415-9 for further details.

applicants who account for the vast majority of applicants. The expansion of the estimated gap suggests that the race coefficients are picking up racial preferences. A similar pattern is observed for the overall rating, where not only does the racial gap between Asian American and African American applicants expand as controls are added, it actually reverses sign. In the sparse model, Asian Americans receive higher overall ratings relative to African Americans, but when all the controls are added, Asian American applicants appear significantly worse.

Examining how the race coefficients change as we alter the set of controls speaks to the relative strength of each group on the observable components in the model. This is useful information, since economists often assume that selection on unobservables works in the same direction as selection on observables. We take this idea one step further by evaluating the average "observed" strength for each racial group across the various Harvard ratings. In particular, we calculate the index of observables (except race and year) for each applicant by multiplying their observable characteristics with the estimated coefficients and summing all the terms. We then subtract off the white mean of the observables index and divide by the variance. Finally, we average within each racial group, labeling this quantity the average index. The average index measures how strong each group is on observable characteristics associated with each rating relative to whites, while the race coefficients measure racial preferences and differences in unobservable characteristics common within racial groups. When the race coefficient and average index move in opposite directions, the case for race playing a role in the rating—as opposed to just proxying for unobserved characteristics—is strengthened. This pattern can only occur if either racial preferences are strong and/or selection on observables works in the opposite direction as selection on unobservables.

The race coefficients and average indices for all of Harvard's ratings are displayed in Appendix Table F8.[A10] In Appendix Figure F3 we graphically display the race coefficients and average index for the academic, extracurricular, overall, and personal ratings. For the academic and extracurricular ratings, the race coefficients have the same signs as the index of observables. This suggests that if we were able to add even more controls, the race effects would likely attenuate to zero. For example, in the case of academics, we exclude information

[A10]The full sets of ratings model coefficients are available in Tables B.6.1R–B.6.4R of Document 415-9.

related to the number of AP exams, AP exam scores, and academic awards.[A11] We know from Table 1 that Asian Americans are stronger on AP exams, and are likely stronger on other unobserved academic measures.[A12]

While the race coefficients in the academic and extracurricular ratings likely reflect selection on unobservables, this is not the case for the overall rating. The race-related coefficients move in the opposite direction of the average index. African American and Hispanic applicants are worse on the average index, while Asian American applicants are stronger. In fact, the ordering of all the race coefficients and average index values is exactly opposite: the order of the race coefficients from largest to smallest is African American, Hispanic, White, and Asian American while the reverse pattern is seen in the average index. A conflicting pattern between the race coefficients and the observable strength of each group is strong evidence of racial preferences. Supporting this interpretation is the fact that Harvard acknowledges that the overall rating incorporates racial preferences.[A13]

The personal rating shows the same pattern as the overall rating. African American and Hispanic applicants receive large bumps in their personal ratings when controlling for all other factors (including all the other ratings other than the overall rating), but are worse on the average index. Asian American applicants, on the other hand, are penalized relative to white applicants but are stronger on the observables that predict the personal rating. Again, the ordering of the race coefficients and average indices by racial group is flipped. This pattern strongly suggests that racial preferences play a large role in the personal rating, similar to the overall rating. As a result, it is incorrect to include the personal and overall ratings in an admissions model focused on estimating racial preferences unless the researcher also calculated the effect of racial preferences through these ratings.

While not displayed in Appendix Figure F3, there is also evidence that Asian American applicants are discriminated against in the school support ratings and alumni personal rating.

[A11] AP exam data is only available for the last two admissions cycles contained in the data.

[A12] The positive coefficient for Asian American applicants in the extracurricular rating model likely reflects differences in the underlying activities that Asian Americans pursue, relative to whites, which are not captured by the model. For example, 45% (27%) of the primary extracurricular activities for white (Asian American) applicants are sports-related, leaving more lines available on the college application for Asian Americans to report non-sports related activities. Note that these numbers include ALDC applicants, see Trial Exhibit DX 680.

[A13] See Document 421-9, pp. 259, 288 and 422.

For each of these ratings, the Asian American coefficient is negative and significant, while Asian American applicants are stronger than white applicants on the index of observables. Since the size of the racial penalty in each of these ratings is substantially smaller than for the personal rating, we take a conservative approach and include these ratings in our preferred admissions model. However, it is important to point out that, by controlling for the school support and alumni personal ratings in all the other ratings models, we are stacking the deck against finding evidence of discrimination. The fact that we still find strong evidence of racial preferences in the overall and personal rating is all the more compelling.

Going beyond selection on observables versus selection on unobservables, there is additional evidence that the personal rating is a tool to implement Harvard's preferences over the composition of their admits. For example, in the personal rating model the interaction between African American and female and African American and disadvantaged are significantly negative, implying racial preferences are muted for these two groups. The share of applicants who are female or disadvantaged is significantly higher for African Americans than for any of the other three major racial/ethnic groups, so if Harvard is interested in balancing within-race characteristics then we would expect to see muted preferences for African American applicants who were female or disadvantaged.[A14] The only other rating that has this pattern is the overall rating, a rating that we know Harvard uses to directly implement preferences.[A15]

[A14] Table 3.1R of Document 415-9 shows descriptive statistics by racial/ethnic group, including share female and share disadvantaged.

[A15] Harvard appears to use the overall and personal ratings to give bonuses for other groups as well. As discussed in Arcidiacono, Kinsler, and Ransom (forthcoming, footnote 27), ordered logit models of the ratings that further include LDC applicants show either no legacy tip or a very small tip except in two cases: legacies receive a significant bonus in both the overall and personal ratings.

D Model Fit

In this section, we describe how information in the public domain helps us more precisely estimate the underlying distribution of the index of applicant observables, AI_i. To pin down the shape of the observable index distribution, we rely on the observed admit rates across deciles of the true AI_i distribution for Asian Americans. These admit rates are presented in Table 9.1 of Document 415-9.

We build off the simple approach described in the text by drawing an initial index for each applicant from a standard normal distribution. Given the initial draw, applicants are sorted into deciles. We then add flexibility by assuming the true underlying distribution of observables is a weighted sum of the initial draw, its square, its square interacted with whether the value was positive, and its exponential. The predicted admit rates for decile k are then calculated as the average of $\frac{\exp(AI_{ik})}{1+\exp(AI_{ik})}$ for all i applicants in decile k. The weights on the various components of the distribution are estimated using the method of simulated moments, matching the predicted admit rates by decile to the observed distribution of admit rates across deciles for Asian American applicants.

Table F9 illustrates that the estimated flexible distribution precisely matches the observed Asian American admit rates across the deciles of the admissions index. For comparison purposes we also show how well a standard normal and log-normal distribution match the data. While the normal distribution does fairly well, the log-normal struggles to match admit rates in the left tail of the distribution.

To calculate the implied R^2 assuming that AI_i is distributed according to the estimated flexible distribution, we complete the following steps.

1. With the underlying distribution of AI_i known, take draws from this distribution (using its parameter estimates obtained from the method of simulated moments estimation described above) and draw ϵ_i's from a logistic distribution.[A16] Assign the highest N_A values of $AI_i + \epsilon_i$ to match the total number of admits to Harvard.

2. Calibrate a logit model of the simulated admissions decisions from the previous step

[A16] Implicit in this step is an assumption that the distribution of AI_i for the population has a similar shape to the distribution of AI_i for Asian Americans.

on AI_i and a constant such that the overall admit rate and the Pseudo R^2 match the actual Harvard data and the admissions model from Document 415-9.

The coefficient on AI_i will be larger (smaller) than one if the variance for Asian Americans is smaller (larger) than the variance for the populations as a whole.

3. Compute the implied R^2 of the model, which is the fraction of variance in $AI_i + \epsilon_i$ explained by AI_i:

$$\begin{aligned} R^2 &= \frac{Var\left(AI_i\right)}{Var\left(AI_i + \epsilon_i\right)} \\ &= \frac{Var\left(AI_i\right)}{Var\left(AI_i\right) + \frac{\pi^2}{3}} \end{aligned}$$

The implied R^2 under the flexible distribution for AI_i is 0.89. It is important to note that this value is a bit misleading. The increase in the R^2 of our flexible distribution relative to the normal distribution comes from the left tail of the distribution: those who have virtually no chance of being admitted.

While the R^2 of the latent index is sensitive to the tails of the distribution, this is not true of accuracy. Assuming AI_i is normally distributed results in an accuracy rate for admits of 64.07%; using the flexible distribution results in an accuracy of 64.09%.[A17] The overall accuracy rate is 96.08% and 96.09%, respectively. Note that an admissions model with no controls would lead to an overall accuracy rate of 90%.[A18]

[A17] The accuracy rate for admits is calculated as the share of the 5.45% of simulated admits based on AI_i and ϵ_i that are in the top 5.45% of the AI_i distribution.

[A18] An admissions model with no controls would randomly assign 5.45% of applicants as admits and 94.55% as rejects. The accuracy rate would then be given by $94.55\% \times 94.55\% + 5.45\% \times 5.45\% = 90.05\%$.

E Difference in the Expert Reports

The results in this paper form the basis of the plaintiff's argument in the *SFFA v. Harvard* lawsuit with regard to Asian American discrimination. Peter Arcidiacono served as expert witness for the plaintiff in the lawsuit, while David Card was the defendant's expert witness. Once the reports of the two experts became public, a set of economists led by Michael Keane filed a brief in support of Arcidiacono's analysis on three key dimensions: (*i*) the exclusion of ALDC applicants from the model, (*ii*) the exclusion of the personal rating from the model, and (*iii*) the interaction of race with disadvantage status.[A19] As we illustrate below—taking everything else Card did at face value—either removing the personal rating or both excluding ALDCs and interacting race with disadvantaged status results in a negative and significant penalty against Asian Americans.

In response to the Keane brief and in support of Card, a number of other prominent economists (led by Sue Dynarski) signed an amicus brief rebutting these points.[A20] The stature of these individuals in the field—which includes two Nobel Prize winners and two of the top economists in President Biden's administration—as well as the stature of Card has lent great weight to their claims, despite the flaws and internal inconsistencies in their arguments.[A21]

The purpose of this appendix is twofold. First, we directly address the differences between our model and Card's model, focusing on the key issues in the Keane and Dynarski briefs. Here we make clear that modeling assumptions employed by Card and supported by the Dynarski brief conflict with standard practices followed by the rest of the field of economics. Second, we illustrate that the only way to generate a finding of no Asian American

[A19] See https://docs.justia.com/cases/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/450. Followup briefs can be found here https://docs.justia.com/cases/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/624 and here https://raw.githack.com/tyleransom/SFFAvHarvard-Docs/master/AmicusBriefs/SFFAappellateAmicusBrief.pdf with additional signatories.

[A20] https://docs.justia.com/cases/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/499. A follow-up brief is available at https://admissionscase.harvard.edu/files/adm-case/files/legal_-_filing_-_200521_-_economists_-_2020.05.21-15_-_brief_for_amici_curiae_professors_of_economics.pdf.

[A21] The *SFFA v. Harvard* lawsuit was also focused on the size of racial preferences in admissions. The desire to preserve those preferences may have served as the rationale for the Dynarski briefs supporting such a flawed approach. However, the Dynarski briefs were focused solely on the Asian-American discrimination claim, not whether Harvard's racial preferences were appropriate.

discrimination using the Harvard data is to make multiple modeling choices that are each at odds with standard approaches.

E.1 Modeling Differences

The key finding from our analysis is that typical Asian American applicants to Harvard are held to a higher standard than their white counterparts. If the average Asian American applicant were treated like a white applicant, their admit rate would increase by 1 percentage point off a baseline admit rate of approximately 5%. Yet, Card is able to generate a finding of no Asian American discrimination. We first outline the three key differences between our model and Card's. We then discuss other differences where there is more room for debate or where the differences are sufficiently small as to not warrant attention. Our findings are robust to all of these other differences in modeling choices.

We now discuss the three key issues that serve as the points of contention in the Keane and Dynarski briefs and explain why the approach by Card is incorrect.

1. **Inclusion of ALDCs:** Card and the Dynarski brief argue that it is not possible to estimate admission preferences for typical applicants without also including in the model recruited athlete, legacy, dean's list, and faculty/staff related applicants (ALDC). From an econometric perspective this is incorrect. Admissions preferences for typical (non-ALDC) applicants can be recovered by estimating a model including only typical applicants, despite the fact that the two groups are competing for the same seats. The strength of the ALDC pool will be reflected in the admission threshold typical applicants need to overcome to be admitted. A more detailed argument is provided in Appendix A. The inconsistency of this argument can be seen in the treatment of foreign applicants by both experts.[A22] Neither expert used foreign applicants in the estimation of their models despite foreign applicants being a part of the admissions process and making up more than 10% of the admitted class each year. Yet, Card did not claim that the admissions models are biased by the exclusion of this applicant group.

[A22] The inconsistency is also revealed by Card's analysis focusing only on female applicants and only on applicants from California. Given that the model that Card uses for his subgroup analysis makes the same faulty assumptions as the one that produced a null effect for Asian Americans as a whole, it is also no surprise that it produced a null effect for these subgroups.

The reason why it is important to exclude ALDC applicants is that preferences work differently for this group. For example, Table 6 in Arcidiacono, Kinsler, and Ransom (forthcoming) shows that academic and extracurricular ratings matter less for ALDC applicants, and thus including them in the model without fully interacting everything distorts these estimates for typical applicants. Rather than interact ALDC with all other attributes, we take the simpler approach and exclude them from the model. Note that more than 97% of Asian American applicants are typical applicants.

Not only does Card include ALDC applicants in his admissions model, he includes them in all of his descriptive analysis. This is misleading since he makes claims about the relative strengths of white and Asian American applicants that are distorted by the presence of ALDC applicants. As an example, in Exhibit 2 of Card's rebuttal report (Document 419-143), he compares how Asian American and white applicants fare on Harvard's academic, extracurricular, personal, and athletic ratings individually as well as the likelihood of obtaining a two or better on at least three ratings. Since recruited athletes receive high athletic ratings and ALDC applicants as a group tend to receive high personal ratings, these comparisons are uninformative about how typical Asian American applicants compare to typical white applicants. This pattern of ALDC inclusion is repeated throughout Card's reports; indeed, Card does not do any analysis where ALDC applicants are taken out.[A23] As a result, much of the supporting evidence Card relies on to argue that Asian American applicants are weaker on non-academic dimensions or *less multidimensional than white applicants* is not germane to our sample of typical applicants.

While we believe our decision to exclude ALDC applicants is appropriate, estimates of the Asian American penalty for typical applicants remain large and statistically significant when ALDC applicants are included in the model. Table 7.2R of Document 415-9 shows that the probability of admission for a typical Asian American applicant increases from 5.2% to 6.1% when treated as a white applicant even when legacy,

[A23]This is in contrast to Arcidiacono, who does his analysis both with and without ALDC applicants. Moreover, Harvard's own Office of Institutional Research (OIR) presents descriptive analysis that excludes legacies and athletes (see Trial Exhibit P009 and Appendix Figure F2).

donor, and faculty/staff applicants (LDC) are included in the model.[A24] Without LDC applicants in the model, the same thought experiment yields an increase in admissions chances from 5.2% to 6.2%. While we have not estimated our preferred admissions model with recruited athletes, Tables B.7.1 and B.7.2 of Document 415-8 illustrate that in a slightly altered admissions model, the addition of ALDC applicants mildly reduces the negative impact of Asian American status for typical applicants.[A25] This is similar to what we find when adding only LDC applicants to our preferred specification.

There are two important takeaways from the admissions models that include ALDC applicants. First, the estimated penalty for typical Asian American applicants remains large and statistically significant. Second, there is no evidence of an Asian American penalty for ALDC applicants. These two results are not in conflict, as there is no ex-ante reason why discrimination should work the same across all groups of applicants. In fact, by belonging to one of the special applicant groups, Asian American applicants may be able to overcome stereotypes that hold typical Asian American applicants back.

The lack of a penalty against Asian American ALDC applicants should not diminish claims that Harvard employs admissions practices that discriminate against Asian Americans. More than 97% of Asian American applicants are not ALDC, meaning that nearly all Asian American applicants face an explicit penalty in admissions. Moreover, the very existence of ALDC preferences works to the detriment of the overwhelming majority of Asian Americans. ALDC applicants are predominantly white, and as we show in Arcidiacono, Kinsler, and Ransom (forthcoming), the elimination of either legacy or athlete preferences would increase the number of Asian American admits by more than 4%.

2. **Inclusion of Personal Rating:** Card and the Dynarski brief argue that Harvard's personal rating is an appropriate control in an admissions model whose purpose is to estimate racial preferences. By Card's own admission, if a control directly incorporates

[A24] When LDC applicants are included, the model also includes indicator variables for legacy, double legacy, faculty or staff child, donor connections, interactions between legacy and race, and interactions between faculty/staff/donor connections and race.

[A25] See Section 8 of Document 415-9 for a detailed discussion of how this model differs from our preferred model.

racial preferences, it is improper (see page 10 of Document 419-141). Below we discuss the overwhelming evidence that racial preferences influence the personal rating.

Descriptive Support

There is a strong positive relationship between academic strength and the personal rating—and indeed all the ratings. See Tables F2, F3, and F4. Yet, despite being strongest on academics, Asian Americans score the worst on the personal rating. To put this in perspective, African Americans in the top decile (10th) of academic strength receive high personal ratings at over twice the rate of African Americans in the 3rd decile; both score better than Asian Americans in the top decile.

The overall rating is explicitly allowed to incorporate racial preferences, and in fact Card excludes the overall rating from his models on these grounds. Yet, the descriptive patterns for race and academic index decile are fundamentally the same for the personal rating and overall rating (see Figure 1 and Table 4). All other ratings show reasonably consistent patterns with regard to race within academic index deciles except these two. One could argue that the personal rating is different in that it could reflect overcoming socioeconomic disadvantage or racism. But Asian Americans are substantially more likely than whites to be classified as disadvantaged (Table 1) and surely would be more likely to experience discrimination than their white counterparts.

Support from Ratings Models

Estimates of models of the personal rating show a large penalty against Asian Americans regardless of the set of controls (see Table F7; for all the intermediate models see Table B.6.3R of Document 415-9). The coefficient on Asian American in the personal rating model is twice the magnitude of any of the other ratings models (see Tables F6 and F7). More importantly, the overall effect on the personal rating is substantial: if Asian Americans were treated as whites their average probability of getting a 2 or better on the personal rating would increase from 17.8% to 21.6%, an over 20% increase.

One concern is that the personal rating model suffers from omitted variable bias. How-

ever, Asian Americans are on average stronger than all racial groups on the observables associated with the personal rating and all other ratings (see Figure F3 and Table F8). Further, Asian Americans are at least as strong as whites on the non-academic observables associated with the personal rating (see Table B.6.13R of Document 415-9). Following Altonji, Elder, and Taber (2005), we would expect any selection on unobservables to move in the same direction as observables, implying that we are likely *underestimating* the penalty Asian Americans receive in the personal rating.

Finally, the personal and overall ratings are the only ones where: (1) the ordering of the race coefficients is the opposite of the strength of the racial groups on the observables (2) the Female x African American coefficient is negative and significant, and (3) the Disadvantaged x African American coefficient is negative and significant (see Tables F6, F7, and F8). These last two patterns are consistent with using the personal and overall ratings to balance within-race gender and disadvantaged status. The fact that these two ratings behave so similarly again suggests that the personal rating incorporates racial preferences since Harvard acknowledges that the overall rating can be a function of race.

Support from Alumni Ratings

While Harvard admissions staff do not typically interview applicants, Harvard alumni do. And here we find a much smaller penalty on the personal rating assigned by alumni who actually meet the applicant, consistent with actually meeting the applicant reducing stereotypes (see Table F7). Additionally, alumni interviews are the place where one might expect discrimination to occur, since alumni have not experienced the same type of training as admissions staff, particularly as it relates to implicit bias. Yet, alumni rate Asian American applicants higher.

Support from UNC

The University of North Carolina (UNC) also uses a personal rating to evaluate applicants. While UNC is not quite as competitive as Harvard, out-of-state admit rates

at UNC are around 13%, making it a highly competitive university. In contrast to Harvard, Asian American applicants to UNC are rated just as well as whites on the personal rating and there is no evidence of Asian American discrimination in admissions. This further illustrates that Harvard is using the personal rating as a means of racial balancing (see footnote 51).

In addition to all of the above, Harvard's own behavior regarding the personal rating suggests that race was a factor admissions staff took into account when assigning a score. During the period we study, Harvard's guidelines for assigning the personal rating made no indication that race could not be used (Trial Exhibit P001). However, in the summer prior to the start of the *SFFA v. Harvard* trial, Harvard altered its guidelines for assigning the personal rating to explicitly state that race was not to be considered (Trial Exhibit P633).

It is important to point out that Card estimates no models of the personal rating that do not show a significant and large negative penalty against Asian Americans.[A26] He also shows no evidence that Asian Americans are worse than white applicants on the observables—as a whole or non-academic—associated with the personal rating. Rather, Card argues that Asian Americans are worse on non-academic factors in his model of *admissions* where preferences for ALDC applicants are *included* as part of the non-academic factors.[A27] Note that excluding the personal rating, but taking Card's approach on every other modeling choice—including having ALDCs in the estimation sample—reveals a negative and significant penalty against Asian Americans.

Arguing against Asian American discrimination in the personal rating despite the con-

[A26] Card does present an analysis where he 'corrects' the personal rating by using the predicted values absent race from Arcidiacono's rating model. This correction is also applied to the academic and extracurricular ratings but to none of the other ratings. Since all the variables that appear in the models of the ratings also appear in the admissions model, there is no exclusion restriction and the model will approximate one where these three ratings are excluded, with the difference being the non-linearities in the controls. Doing this—along with Card's other assumptions including keeping ALDCs in the sample—does not show a significant penalty. However, removing *all* the ratings does show a penalty against Asian Americans. See column (4) of Tables B.7.1R and B.7.2R in Document 415-9.

[A27] See Document 419-143 (p. 30) and the trial judge's ruling (Document 672, p. 58) where these results were cited.

sistently large, negative, and significant effect of being Asian American in all models of the personal rating and despite Asian Americans being stronger on the observables associated with the personal rating is disheartening. Further, such an argument sets a dangerous precedent by providing a blueprint for how to discriminate against other groups. An institution can simply create a rating—say, likability—and then avoid an accusation of discrimination by claiming that it is not discriminating; rather it is just that members of the discriminated group are not likable.

As we discuss in Section 5.1, adding the personal rating to our preferred admissions model cuts the Asian American penalty by less than half. Even if one believes the personal rating to be an appropriate control (despite the overwhelming evidence to the contrary), there is still evidence of a large and statistically significant penalty for typical Asian Americans. We believe a more reasonable interpretation is that bias in the personal rating accounts for a little less than half of the Asian American admissions penalty.

3. **Exclusion of Race-Disadvantaged Interactions:** Card and the Dynarski brief argue that it is improper to include interactions between race and disadvantaged status in the admissions model, believing there is no theoretical basis for doing so.[A28] However, including these interactions in the model reveals that disadvantaged status matters quite differently for African American and Hispanic applicants. This is especially true for African Americans who receive no bump for being disadvantaged. Card himself writes, "The typical approach ... would be to include an interaction between race and disadvantaged status only if the effect of being disadvantaged is different for Asian-American and White applicants (or, equivalently, if the effect of race is different for disadvantaged and non-disadvantaged applicants)" (see page 49 of Document 419-141).

   Including interactions between disadvantaged status and race matters for estimating the magnitude of Asian American discrimination. Because African American and Hispanic applicants receive smaller bumps for being disadvantaged, excluding the inter-

[A28] That there is no theoretical basis for doing so is false. Indeed, Harvard's own Office of Institutional Research (OIR) estimated admissions models that included these interactions (Trial Exhibit P009 and Trial Exhibit P028). See Arcidiacono (2005) for a similar approach.

actions dilutes the overall impact of disadvantaged status on admissions. This results in a smaller estimated Asian American penalty since Asian American applicants are significantly more likely to be disadvantaged relative to white applicants.

On these three important modeling differences, the positions taken by Card (and supported by the Dynarski briefs) are simply not defensible for evaluating how and whether discrimination occurs against Asian American applicants. The fact that these modeling choices are needed to defend Harvard provides strong supporting evidence of the strength of the Asian American discrimination finding.

There are other differences between the Arcidiacono and Card models. Some are minor and others are not as clear as the three issues above. We go through each of these points here, explain why we made the modeling choices we did, and note that none of them affect the finding of an Asian American penalty in Harvard admissions.

1. **Inclusion of Staff Interview:** Card argues for the inclusion of an indicator for receiving a staff interview in the admissions model. This indicator is basically irrelevant when ALDC applicants are excluded, as only 1.3% of typical applicants receive an interview. However, over 20% of ALDC applicants receive an interview. Thus, access to obtaining a staff interview is clearly a function of applicant status, and will likely relate to other attributes, including race. This is also consistent with how Card treats the variable, choosing to include an indicator for obtaining an interview instead of the resulting applicant rating generated by the staff member. It is also relevant to Card's claims that whites are stronger than Asian Americans on nonacademic characteristics as this is one of the controls included.[A29]

2. **Inclusion of parental occupation and intended career:** Card argues for the inclusion of parental occupation and intended career in the admissions model. In principle this is a reasonable point. We do not include these variables in our preferred model because how they are coded is wildly inconsistent across admissions cycles, forcing the analyst to make a number of *ad hoc* choices regarding how to code and combine

[A29] While such a low rate among typical applicants would make little difference to the *average* characteristics of applicants, it is relevant for the share in the top 10% of applicants which is where the claim is made.

various groupings to generate a consistent measure. This, coupled with the fact that there are numerous other variables already included in the model that capture parental background and the interests of the applicants, is why we exclude these variables in our preferred model.

To illustrate, in the database made available as part of *SFFA v. Harvard*, parental occupation is available through the Common Application using either a Bureau of Labor Statistics (BLS) code or a Common Application code. The use of the two codes varies across cycles and the categories within each occupation code change over time. Appendix Table F10 provides evidence on the inconsistency of the parental occupation variables. For example, between the Classes of 2014 and 2015, the number of fathers in the 'Other' occupation classification nearly triples from 1,593 to 4,608. Between 2017 and 2018, the number of fathers who are unemployed drops from 1,300 to 5.

In addition to the lack of consistency, there is little evidence that parental occupation matters beyond helping to determine whether an applicant is disadvantaged. While parental occupation is included on an applicant's summary sheet, it does not appear to be an important part of the evaluation process. For example, the reader guidelines for 2017 (Trial Exhibit DX 016), 2018 (Trial Exhibit P001), 2019 (Trial Exhibit P071), and 2023 (Trial Exhibit P633) never discuss parental occupation, but do discuss scores, ratings, interviews, GPA, disadvantaged status, etc.[A30] Additional evidence on the inconsequential impact of parental occupation comes from Trial Exhibit DX 024, a discussion guide for the 2012 casebook. This guide walks readers through 12 pseudo applications and discusses key features of each application and admissions outcomes. Across the 12 applicants and 12 pages of discussion, parental occupation is never discussed beyond one mention of a parent being blue-collar.[A31] For these reasons,

[A30] Further evidence that parental occupation is unimportant comes from Trial Exhibit P238. This document shows an internal email conversation among Harvard employees early in the admissions cycle for the Class of 2017. The following is a direct quote from the email, "RMW just noticed that parent2 employer field not showing up on the reader sheets. Turns out I had cut it by accident...Though if they're only just noticing this now, I do wonder how important it is or how carefully they're paying attention."

[A31] Parental occupation is discussed in deposition testimony as a tool to infer disadvantaged status. See p. 201 of Document 421-9 ("Q. How does Harvard determine whether or not an applicant is socioeconomically disadvantaged? A. ...We also have information at the outset about the parents' educational and professional backgrounds."); p. 59 of the deposition of Christopher Looby ("Q. What types of information would you

we exclude parental occupation from our preferred model. For similar reasons, we exclude an applicant's intended career. This is a variable that varies considerably across admissions cycles and—since we already account for intended major—seems unnecessary.[A32]

Card advocates including both of these variables (Document 419-141). For the occupation controls, Card harmonizes the reported parental occupation codes by mapping Common Application codes to major and minor groups in the BLS-Standard Occupational Classification System. Major and minor groups are then combined into broad occupational categories. There are 24 occupational classifications for mothers and fathers, with little explanation for the chosen groupings. For example, business executive, business and financial operations, and other management are included as separate categories. Low skill is separate from construction and protective service. Further evidence that the occupation variables are not especially informative is that the second most common occupation among both mothers and fathers is "Other" (see pp. 178–179 of Document 419-141).

While we believe the occupation category and intended career variables are unreliable and superfluous, we test the robustness of our preferred model to their inclusion.[A33] When occupation and intended career are added to the preferred model, the average marginal effect associated with being Asian American is -0.75% and statistically significant at the 5% level.[A34] While smaller than our preferred model estimates, it still indicates a large penalty for Asian American applicants relative to white applicants.

3. **Pooled vs Yearly Model:** Card and those who signed the Dynarski amicus brief argue that the appropriate approach for estimating Harvard admission preferences is to estimate separate models for each admission cycle (six in all). We disagree and argue that a pooled model with appropriate interactions between admission cycle and

assess in trying to determine whether you should code an applicant as disadvantaged? ... A. Could be parent jobs.") [Document 419-143, fn. 56].

[A32]See Appendix Table F11 for information on how intended career varies by admissions cycle. Intended career is also never discussed in the reader guidelines made public over the course of the trial: 2017 (Trial Exhibit DX 016), 2018 (Trial Exhibit P001), 2019 (Trial Exhibit P071) or 2023 (Trial Exhibit P633).

[A33]Adding these 64 variables to the model increases the total number of controls by more than 15%.

[A34]See Table 8.2N in Document 415-9.

applicant characteristics is superior. In particular, we include interactions between admission cycle and applicant characteristics such as gender, disadvantaged status, and intended major. These interactions allow Harvard to balance their class along these dimensions each cycle. However, there is no reason to believe that Harvard values test scores, high school GPA, and profile ratings differently each year. As the next section will show, this disagreement has very little impact on the results. If we estimate Card's yearly models under a more reasonable set of assumptions as outlined above, a large and statistically significant Asian American penalty emerges.

There are other minor differences between Card's model and ours, but they have little impact on the findings and have received little attention from the signatories of either side's amicus brief. In the next section, we turn to Card's model and show the fragility of his no discrimination finding.

Before jumping to Card's model, it is also important to point out that Harvard's Office of Institutional Research (OIR) reports admissions model estimates in Trial Exhibit P009 and Trial Exhibit P028 that pool application cycles, include both ALDC applicants and the personal rating, but also include race-by-disadvantaged status interactions. In each of these models, there is a statistically significant Asian American penalty.[A35]

E.2 The Fragility of a No Discrimination Finding

In this section we start from Card's baseline specification where he finds an insignificant Asian American penalty, and explore how simple and reasonable alterations from this baseline lead to changes. The broader point is that, while our preferred specification is quite robust, Card's specification that finds no penalty is very fragile. In all of the analysis below we focus on modeling admissions for typical applicants since this is the appropriate sample to study.

[A35] The OIR models respectively cover the Classes of 2007–2016 (Trial Exhibit P009) and 2009–2016 (Trial Exhibit P028). Both sets of models also include foreign applicants, which are completely excluded from the expert reports on both sides of the *SFFA v. Harvard* case.

E.2.1 Pooled Models

We begin by exploring the sensitivity of a pooled admissions model proposed by Card that is capable of generating a small and insignificant Asian American penalty. This is not Card's preferred model, but is a good starting point for understanding the importance of the assumptions related to the controls included in the model. For full details of the model, see Section 5 of Document 419-141. Again, we focus on versions of this model that exclude ALDC applicants, but it is important to note that Card never estimates an admissions model excluding this special set. As we illustrated earlier in Appendix A, it is inappropriate to include these applicants unless indicators for ALDC are interacted with all the other applicant attributes. A simpler approach is to just exclude them.

The key differences between Card's pooled model and our preferred pooled specification are: (*i*) inclusion of the personal rating; (*ii*) exclusion of interactions between race and disadvantaged status; and (*iii*) inclusion of parental occupation.[A36] Appendix Table F12 shows how Card's estimated Asian American penalty is affected by the modeling choices associated with points (*i*)–(*iii*). The first row shows that it is possible to construct a pooled admissions model that yields no statistically significant Asian American penalty. The remaining rows show that changing any of the three questionable modeling choices results in a statistically significant Asian American penalty.[A37] Moreover, altering all three components essentially leads to a result that is almost identical to our preferred specification. Thus, the other differences between our preferred model and Card's pooled model have a relatively minor impact.

[A36] While these three differences will be our primary focus, there are other differences between the models. See Document 415-9 for a discussion.

[A37] As discussed in the previous section, the inclusion of race by disadvantaged status is necessitated by the differential effect disadvantaged status has for African American and Hispanic applicants. To see this, rather than interact race with disadvantaged status, we estimate two alternative models where we: 1) include only white and Asian American applicants, and 2) include only non-disadvantaged applicants. In both cases the estimated Asian American penalty is statistically significant and slightly larger than when we interact race and disadvantaged status using the full sample (see rows 3 and 4 of Table F12).

E.2.2 Yearly Models

While Card estimates a pooled admissions model (Document 419-141), his preferred approach is one that estimates admissions preferences separately by year.[A38] The structure of the yearly models is essentially identical to the pooled model in terms of included controls. The benefit of the yearly approach is it allows for variability in the impact of applicant attributes over time. The cost is reduced statistical power and the potential for model overfitting. Approximately 2,000 applicants are admitted each year, and the yearly models will contain well over 200 variables each. In contrast, the pooled model includes approximately 350 variables, but there are more than 11,000 admits across all cycles.

One important difference between Card's pooled and yearly models is the inclusion of total work hours and indicators for an applicant's primary extracurricular activities. The reason these variables are excluded from the pooled model is that they are only available for applicants to the Classes of 2017–2019.[A39] However, the decision to use the detailed extracurricular activities in this particular manner is odd. Data on extracurricular activities come from applicants listing each activity they participated in, the years in which they participated in this activity, the hours per week and weeks per year they participated in the activity, and whether their participation was during the school year or outside the school year. Each of the activities is assigned to one of 29 categories (e.g., work, academics, musical instruments). Card defines a primary activity as an activity the applicant lists in the first or second activity field of her application (Document 419-141). Additionally, the primary activities are collapsed into one of twelve groups in a somewhat arbitrary manner. More

[A38] In this section, we focus on the fragility of Card's yearly models in his initial report (Document 419-141), not his yearly models in his rebuttal report (Document 419-143). The reason for this is that we were able to analyze the models in the initial report as part of a response, but we were not given the opportunity to respond to Card's rebuttal report. While Card's yearly models differ slightly across the two reports, all of his preferred models maintain the same faulty assumptions regarding the personal rating, race-disadvantaged interactions, and parental occupation. The other differences are less relevant for estimating the Asian American penalty.

[A39] Similarly, AP exam scores are only available in the final two admission cycles. However, Card does not utilize these variables in his yearly regressions when they are available, despite the fact that admission is positively correlated with AP performance and Asian Americans take more AP exams and score higher conditional on taking the exams (see Table B.3.1 of Document 415-9). By choosing to exclude AP exams and scores, Card is biasing the Asian American penalty towards zero. Note their exclusion in earlier admissions cycles is not a choice, but is still likely to result in an estimated Asian American penalty that is biased towards zero.

importantly, the level of participation of the activity is done only for the work category, where total work hours are calculated over the course of the applicant's high school career. This distorts the analysis in two ways. First, it overemphasizes the weight that work is given in the process, as work activities are only the eighth most popular activity listed for whites.[A40] Second, white applicants work significantly more hours than Asian American applicants. Yet there are many activities where Asian American applicants invest substantially more hours than white applicants.

As a result, when we investigate the robustness of the yearly models, we consider two cases. First, we look at the case where we take the extracurricular activities defined by Card at face value. Second, we define our own set of extracurricular controls. We use the original 29 activity categories when constructing indicators for each of the first two listed activities. Instead of using the total hours of work over the course of the applicant's high school career, we consider broader groupings of categories and measure participation both by counting the number of grades in which the applicant participated in each activity and indicating whether the applicant's total accumulated hours in a category was above the median for those who had any positive hours in the category. Making these adjustments more precisely accounts for the impact of extracurricular activities on admissions decisions.

Similar to the pooled model robustness exercise, we are interested in whether a finding of an insignificant Asian American penalty using Card's baseline yearly model is robust to: (*i*) inclusion of the personal rating; (*ii*) the exclusion of interactions between race and disadvantaged status; and (*iii*) inclusion of parental occupation. Appendix Table F13 indicates that the finding of no Asian American penalty is not robust. In the first column we present the estimated Asian American penalty when we employ the extracurricular variables as constructed by Card. The marginal effects reported are a weighted average of the year-specific estimates. We find that the magnitude of the estimated penalty in the yearly model is similar to the pooled model when we interact race and disadvantage. However, the result is not statistically significant. Excluding the personal rating or parental occupation leads to a large and statistically significant Asian American penalty. In the second column, we estimate Card's yearly model, but use the corrected extracurricular measures. In this case,

[A40] See Document 419-141 Appendix D, Exhibit 66.

the Asian American penalty is statistically significant when any of (*i*)–(*iii*) are addressed. The final column of the table are the results from the pooled specification and show that the estimated magnitude of the penalty is largely unaffected by moving to the yearly model.

The weighted averages reported in Appendix Table F13 mask important heterogeneity in the size and significance of the Asian American penalty across admissions cycles. In Appendix Table F14 we provide the year-by-year estimates of the Asian American penalty for Card's baseline specification, as well as the robustness checks related to disadvantaged status, the personal rating, and parental occupation. In all models we use the extracurricular variables as defined by Card. For every specification, the estimated penalty is negative in all years except 2019. This pattern is interesting since this is the only admissions cycle to occur after the SFFA lawsuit was filed. The final row of the table reports the average marginal effect across admissions cycles excluding 2019. Here we find that, even when we add to the baseline model race interacted with disadvantaged status, the Asian American penalty is large and statistically significant. When we make all the model adjustments and exclude 2019, the Asian American penalty is 20% larger than in the corresponding yearly specification including 2019 (-0.90 from row (5) from Table F13).

This section has shown that being able to find no significant Asian American penalty among typical applicants to Harvard requires making a number of questionable modeling choices. If any of these decisions are reversed, a statistically significant Asian American penalty appears. This lack of robustness is in sharp contrast to our preferred specification, where altering many of the modeling choices does not alter the main finding.

Online Appendix References

Arcidiacono, Peter. 2005. "Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica* 73 (5):1477–1524. 16, 25

Arcidiacono, Peter, Josh Kinsler, and Tyler Ransom. 2019. "Recruit to Reject? Harvard and African American Applicants." Working paper, Duke University.

———. forthcoming. "Legacy and Athlete Preferences at Harvard." *Journal of Labor Economics* . 2, 8, 12, 15, 23, 24, A16

Document 415-8. 2017. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://docs.justia.com/cases/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/415/1.html. Plaintiff Expert Witness Opening Report. 7, 9, 22, 24, 25, A16, A31

Document 415-9. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://docs.justia.com/cases/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/415/2.html. Plaintiff Expert Witness Rebuttal Report. 6, 12, 16, 19, 21, 23, 24, 26, 35, 36, 37, 38, 39, 40, 41, A15, A16, A17, A18, A31, A35

Document 419-141. 2017. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://projects.iq.harvard.edu/files/diverse-education/files/expert_report_-_2017-12-15_dr._david_card_expert_report_updated_confid_desigs_redacted.pdf. Defendant Expert Witness Opening Report. 5, A31

Document 419-143. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://projects.iq.harvard.edu/files/diverse-education/files/expert_report_rebuttal_as_filed_d._mass._14-cv-14176_dckt_000419_037_filed_2018-06-15.pdf. Defendant Expert Witness Rebuttal Report. 22, A31

Document 421-9. 2017. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://www.courtlistener.com/recap/gov.uscourts.mad.165519/gov.uscourts.mad.165519.421.9_7_1.pdf. Deposition of William Fitzsimmons. 13, 22, A31

Document 672. 2019. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://www.courtlistener.com/recap/gov.uscourts.mad.165519/gov.uscourts.mad.165519.672.0_2.pdf. Findings of Fact and Conclusion of Law.

Trial Exhibit DX 016. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.

com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/D016.pdf. Harvard Reading Procedures for Class of 2017.

Trial Exhibit DX 024. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/D024.pdf. 2012 Casebook Discussion Guide. A31

Trial Exhibit DX 680. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/D680.pdf. Primary extracurricular activity by race. A31

Trial Exhibit P001. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P001.pdf. Harvard Reading Procedures for Class of 2018. 7, 8, 18, 22, A19, A31

Trial Exhibit P009. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P009.pdf. Office of Institutional Research report, "Admissions Part II: Subtitle". 4, A31

Trial Exhibit P028. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P028.pdf. Office of Institutional Research report, "Demographics of Harvard College Applicants". 4, A31

Trial Exhibit P071. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P071.pdf. Harvard Reading Procedures for Class of 2019.

Trial Exhibit P238. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P238.pdf. Email correspondence between admissions office personnel.

Trial Exhibit P633. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P633.pdf. Harvard Reading Procedures for Class of 2023. 7, A19, A31

F Supporting Figures and Tables

Table F1: List of *SFFA v. Harvard* Legal Documents Used

Document	Description
Document 415-8	Plaintiff's expert witness opening report
Document 415-9	Plaintiff's expert witness rebuttal report
Document 419-141	Defendant's expert witness opening report
Document 419-143	Defendant's expert witness rebuttal report
Document 419-1	Deposition of Harvard Admissions Director Marlyn McGrath
Document 421-9	Deposition of Harvard Admissions Dean William Fitzsimmons
Trial Exhibit DX 016	Class of 2017 application reading procedures
Trial Exhibit DX 024	Class of 2012 casebook discussion guide
Trial Exhibit DX 042	Demographic breakdown of applicants, admits and matriculants
Trial Exhibit DX 680	Table of primary extracurricular activities by race
Trial Exhibit P001	Class of 2018 application reading procedures
Trial Exhibit P009	Harvard Office of Institutional Research (OIR) report
Trial Exhibit P028	Harvard OIR report on admissions
Trial Exhibit P071	Class of 2019 application reading procedures
Trial Exhibit P164	"One-pager" for Class of 2018
Trial Exhibit P238	Email correspondence between admissions office personnel
Trial Exhibit P555	Office for Civil Rights Report (1990)
Trial Exhibit P621	Ratings frequencies for baseline sample
Trial Exhibit P633	Class of 2023 application reading procedures
Day 4 Trial Transcript	Transcript of Day 4 of trial
Day 14 Trial Transcript	Transcript of Day 14 of trial
Document 672	Trial court judge's ruling

Table F2: Share Receiving a 2 or Better on Academic and Extracurricular Ratings by Academic Index Decile and Race

Decile	White	African American	Hispanic	Asian American
Panel A: Academic Rating				
1	0.11	0.02	0.03	0.00
2	0.41	0.08	0.05	0.54
3	1.91	0.96	0.68	1.36
4	9.14	6.07	4.45	7.98
5	26.26	23.08	17.04	26.36
6	50.19	48.43	43.83	51.08
7	68.37	68.54	64.28	71.46
8	82.73	80.37	79.63	86.16
9	93.30	93.37	91.47	95.12
10	97.16	94.70	95.26	98.08
Average	45.32	9.18	16.75	60.21
Panel B: Extracurricular Rating				
1	11.41	9.02	9.27	12.97
2	16.35	13.75	12.73	15.99
3	20.14	18.86	15.86	18.57
4	22.02	23.27	18.74	21.59
5	23.83	22.85	20.65	23.67
6	25.08	26.38	23.31	25.51
7	26.64	27.42	27.61	28.34
8	27.31	27.91	24.63	29.78
9	30.45	32.65	28.94	34.92
10	33.04	38.64	29.21	37.98
Average	24.38	15.56	16.84	28.27

Source: Authors' calculations from data presented in Table 5.4R of Document 415-9. Data restricted to non-ALDC applicants from the Classes of 2014–2019.

Notes: Portions of this table also appear in Arcidiacono, Kinsler, and Ransom (2019) as Table 5.

Table F3: Share Receiving a 2 or Better on Personal and Alumni Personal Ratings by Academic Index Decile and Race

Decile	White	African American	Hispanic	Asian American
Panel A: Personal Rating				
1	8.11	9.49	8.48	8.01
2	12.58	15.75	13.16	12.91
3	16.25	23.35	17.77	13.46
4	18.62	28.95	20.39	14.24
5	20.40	33.89	25.60	15.69
6	22.72	35.04	28.41	16.46
7	22.59	40.00	30.03	18.11
8	26.10	39.57	32.20	17.93
9	28.23	40.31	30.24	20.87
10	29.62	46.97	34.21	22.20
Average	21.29	19.01	18.69	17.65
Panel B: Alumni Personal Rating				
1	26.33	30.96	26.29	28.13
2	33.72	39.83	33.42	32.03
3	39.77	46.84	38.59	36.35
4	44.27	55.56	43.86	40.66
5	48.43	59.98	50.32	44.24
6	51.84	62.20	54.50	46.96
7	54.08	69.89	56.90	51.93
8	58.20	67.48	62.44	53.78
9	62.20	70.92	62.89	57.46
10	64.98	73.48	71.05	63.61
Average	49.79	42.79	41.25	50.21

Source: Authors' calculations from data presented in Table 5.6R of Document 415-9. Those with missing ratings are excluded from the calculations. Data restricted to non-ALDC applicants from the Classes of 2014–2019.

Notes: Portions of this table also appear in Arcidiacono, Kinsler, and Ransom (2019) as Table 5.

Table F4: Share Receiving a 2 or Better on School Support Ratings by Academic Index Decile and Race

Decile	White	African American	Hispanic	Asian American
Panel A: Teacher 1 Rating				
1	7.76	7.75	8.85	7.41
2	13.42	13.97	13.87	14.18
3	19.00	19.38	20.03	16.98
4	23.87	25.06	23.60	21.03
5	26.39	29.65	30.19	23.00
6	32.41	36.42	31.94	26.59
7	34.64	40.22	35.62	30.22
8	39.72	46.63	37.68	33.09
9	44.92	47.45	43.60	39.73
10	50.17	55.30	49.47	46.64
Average	30.46	17.15	21.60	30.84
Panel B: Teacher 2 Rating				
1	6.20	5.46	6.42	6.55
2	10.24	11.50	11.00	11.69
3	15.46	16.98	17.77	13.80
4	21.21	22.41	20.81	18.01
5	23.31	31.55	25.54	20.26
6	27.53	35.43	28.97	24.29
7	31.04	35.06	32.77	26.18
8	36.66	39.88	37.32	29.67
9	41.47	42.86	38.59	36.15
10	47.11	50.76	49.74	41.90
Average	27.16	14.83	18.86	27.44
Panel C: Counselor Rating				
1	4.64	4.88	5.72	5.76
2	8.99	10.86	10.15	9.19
3	14.49	16.72	14.83	12.25
4	18.49	20.31	17.32	14.93
5	22.06	26.42	21.06	17.84
6	25.59	32.87	25.26	22.61
7	29.24	35.73	30.35	24.96
8	34.39	38.04	34.15	27.69
9	39.16	43.88	34.32	33.88
10	44.63	49.24	45.00	38.34
Average	25.29	13.86	16.49	25.16

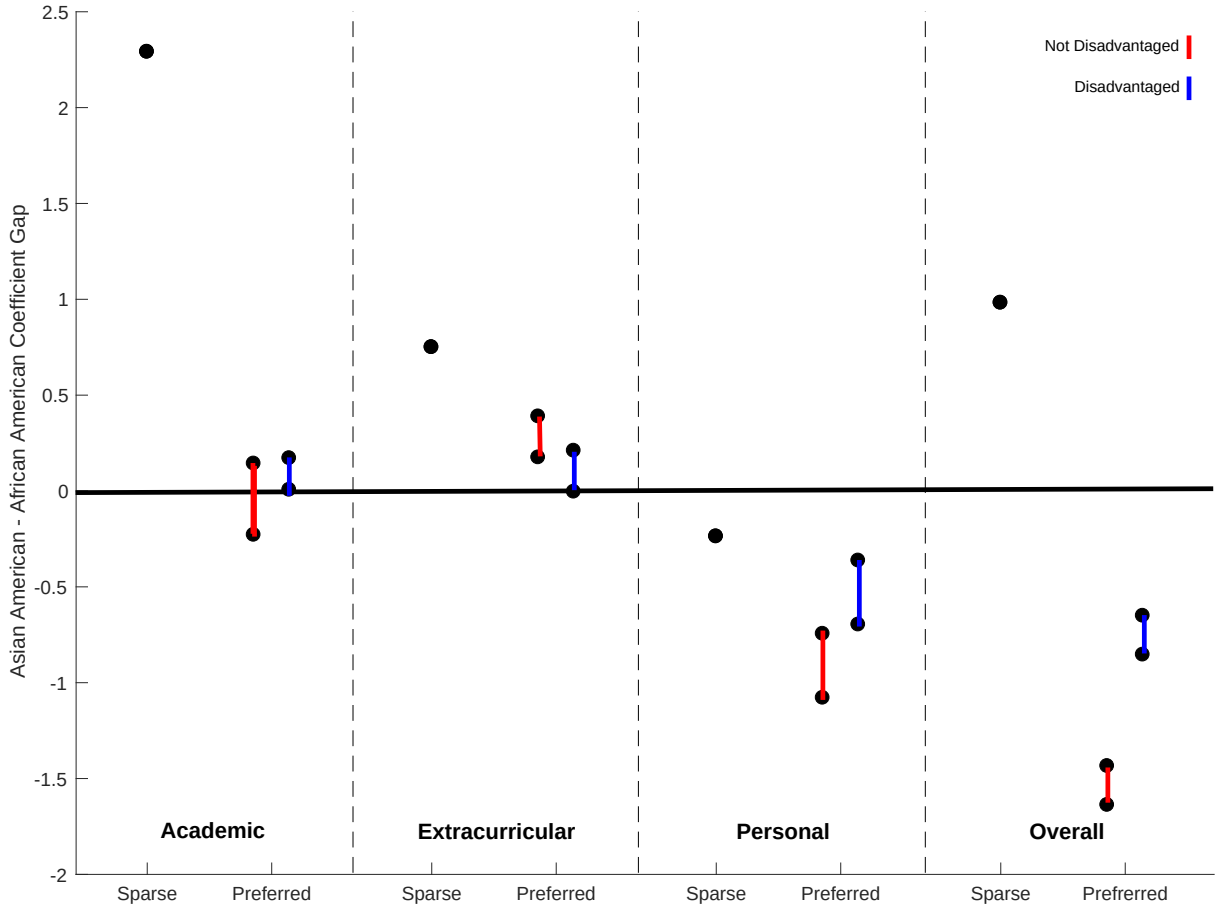
Source: Authors' calculations from data presented in Table 5.5R of Document 415-9. Those with missing ratings are excluded from the calculations. Data restricted to non-ALDC applicants from the Classes of 2014–2019.

Table F5: Admission Rates of Applicants by LDC Status, Race, and Academic Index Decile

	White		African American		Hispanic		Asian American		Total	
Decile	Typical	LDC	Typical	LDC	Typical	LDC	Typical	LDC	Typical	LDC
1	0.00	6.32	0.03	3.19	0.00	0.00	0.00	6.29	0.01	5.27
2	0.39	12.20	1.03	6.61	0.32	11.54	0.20	7.16	0.53	10.47
3	0.56	16.67	5.19	25.36	1.95	8.15	0.64	11.53	1.65	15.56
4	1.82	22.62	12.76	39.94	5.50	30.20	0.86	23.90	3.29	23.72
5	2.57	26.18	22.41	48.92	9.13	42.45	1.86	21.28	4.40	28.39
6	4.20	31.85	29.72	54.73	13.65	41.46	2.49	29.78	5.64	33.70
7	4.79	36.04	41.12	82.43	17.28	48.49	3.98	40.45	6.61	38.51
8	7.53	47.49	44.48	75.01	22.93	49.85	5.12	53.17	8.22	47.66
9	10.77	56.94	54.59	99.90	26.16	43.98	7.55	56.45	10.40	56.67
10	15.27	57.07	56.06	83.43	31.32	95.10	12.69	63.02	14.58	60.64
Total	4.90	33.47	7.58	27.52	6.16	34.73	5.14	36.75	5.46	33.73

Source: Authors' calculations from data presented in Tables 5.1R, 5.2R, B.5.1R and B.5.2R of Document 415-9.

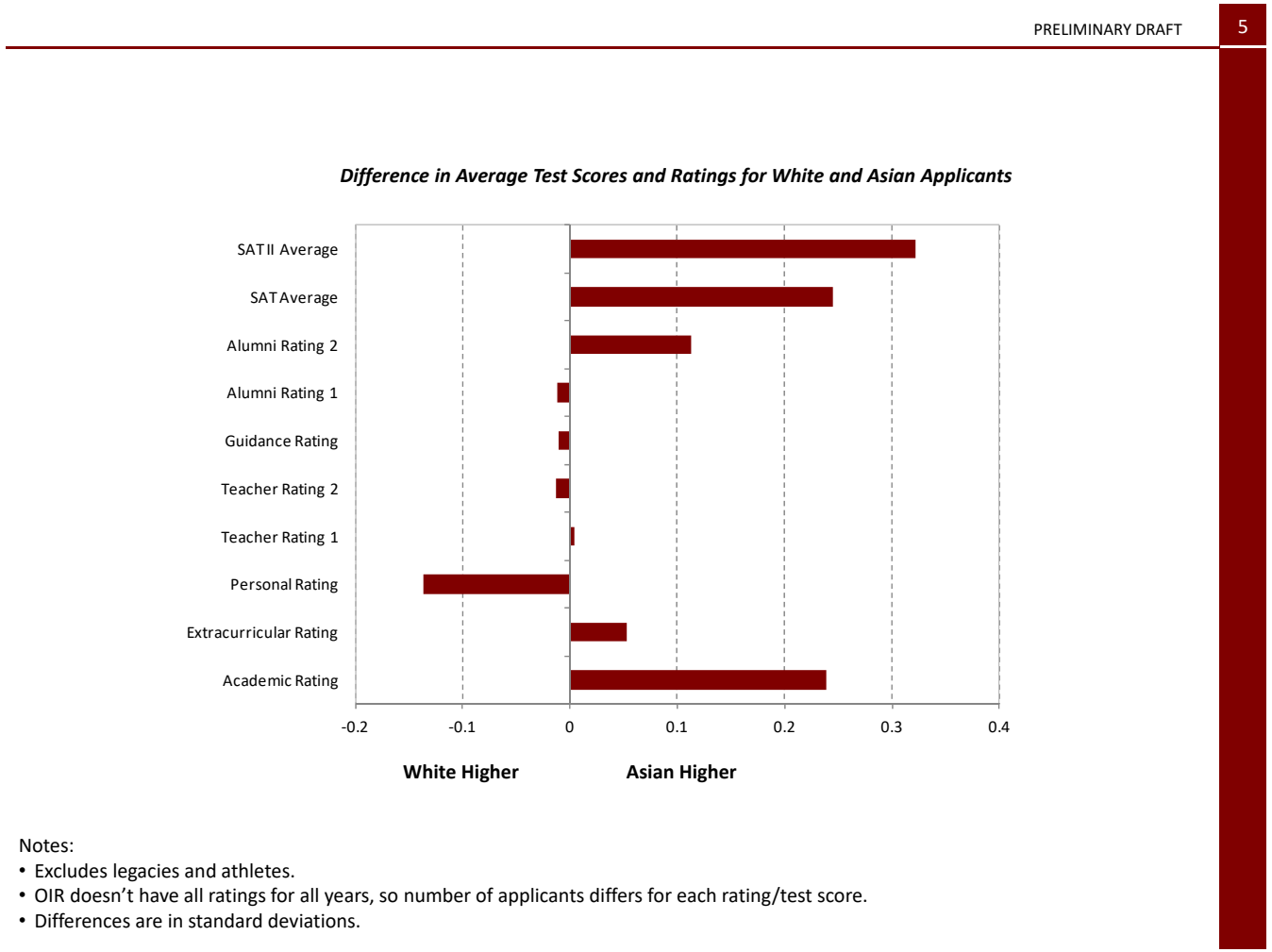
Figure F1: Estimated Ratings Gaps between Asian Americans and African Americans with Varying Number of Controls



Source: Authors' calculations from results reported in Appendix Tables F6 and F7.

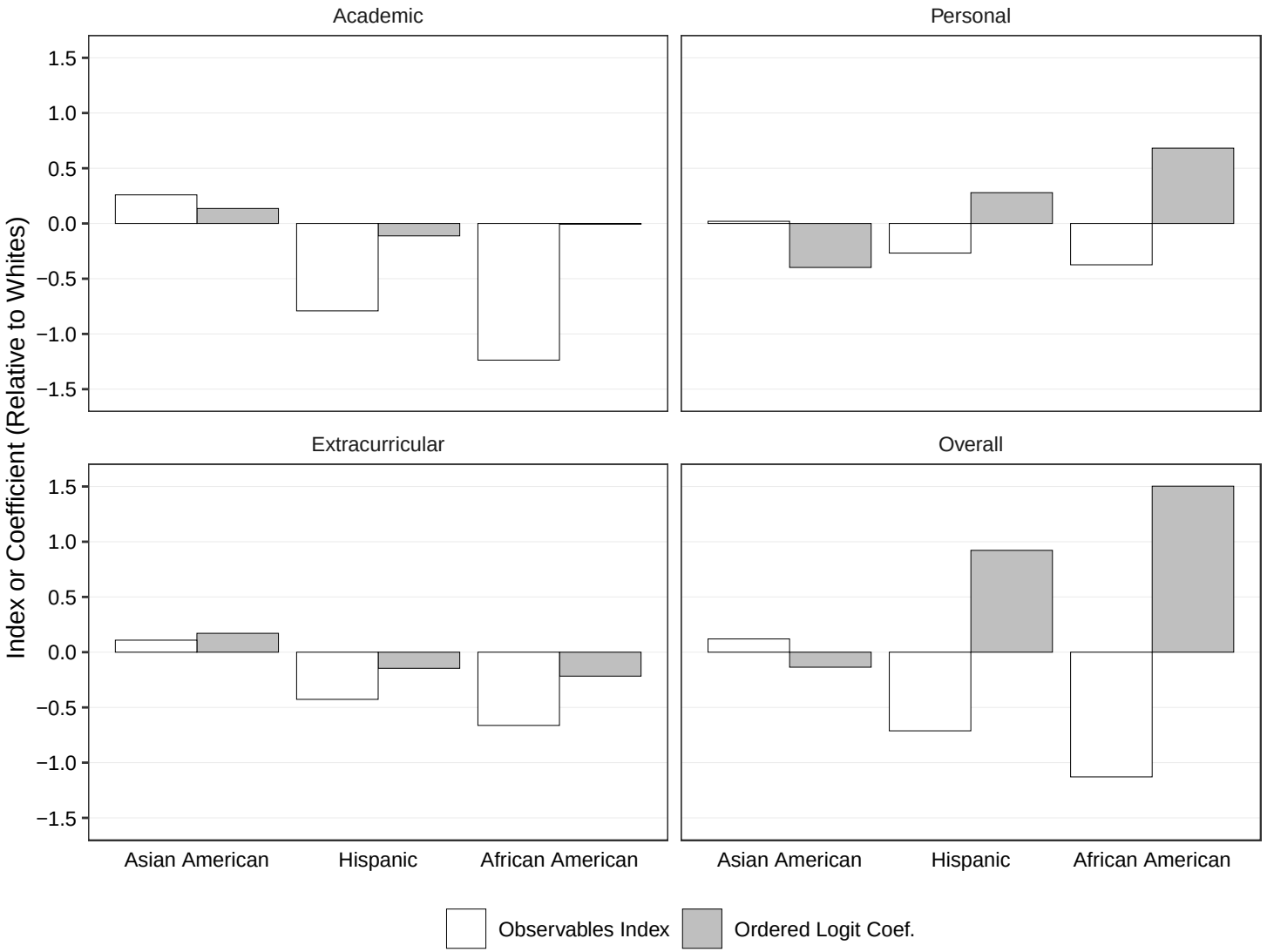
Notes: Dots indicate coefficient estimates for a given rating and specification. Intervals represent the range between men and women by disadvantaged status. "Sparse" refers to a model with relatively few covariates (i.e. Model 1 in the results of Document 415-9); "Preferred" means the preferred model (i.e. Model 5 in the results of Document 415-9).

Figure F2: Harvard OIR Analysis of White and Asian American Applicants



Source: Page 5 of Trial Exhibit P009 (report of Harvard's Office of Institutional Research).

Figure F3: Race Coefficients and Observable Indices by Harvard Ratings



Source: Authors' calculations from Table B.6.11R of Document 415-9.

Notes: "Observables Index" refers to the predicted linear index of observables (i.e. $X_i\hat{\gamma}^R$) after removing race and year effects. "Ordered Logit Coef." refers to the coefficient on race from the ordered logit model of the given Harvard rating.

Table F6: Academic, Extracurricular, and School Support Ratings, Selected Coefficients

	Academic		Extracurricular		Teacher 1		Teacher 2		Counselor	
	Sparse	Preferred	Sparse	Preferred	Sparse	Preferred	Sparse	Preferred	Sparse	Preferred
African American	-1.685	-0.006	-0.503	-0.217	-0.606	0.012	-0.551	0.104	-0.577	0.164
	(0.019)	(0.043)	(0.023)	(0.044)	(0.024)	(0.048)	(0.026)	(0.051)	(0.026)	(0.052)
Hispanic	-0.944	-0.112	-0.302	-0.146	-0.289	-0.023	-0.256	0.024	-0.289	0.017
	(0.017)	(0.037)	(0.021)	(0.036)	(0.021)	(0.037)	(0.023)	(0.039)	(0.023)	(0.040)
Asian American	0.614	0.136	0.246	0.171	-0.048	-0.159	-0.086	-0.203	-0.054	-0.095
	(0.014)	(0.031)	(0.015)	(0.026)	(0.015)	(0.026)	(0.016)	(0.028)	(0.016)	(0.028)
Missing	0.318	0.082	0.133	0.077	-0.015	-0.080	-0.063	-0.115	-0.048	-0.116
	(0.023)	(0.051)	(0.025)	(0.043)	(0.025)	(0.043)	(0.026)	(0.046)	(0.027)	(0.047)
Female	-0.272	0.116	0.207	0.021	-0.001	0.093	-0.027	0.085	0.032	0.034
	(0.011)	(0.034)	(0.012)	(0.031)	(0.012)	(0.032)	(0.013)	(0.035)	(0.013)	(0.035)
Disadvantaged	0.131	0.048	0.372	0.202	0.430	0.188	0.453	0.278	0.451	0.168
	(0.020)	(0.046)	(0.024)	(0.045)	(0.024)	(0.045)	(0.026)	(0.048)	(0.026)	(0.049)
Female X African American		0.097		0.216		-0.068		-0.096		-0.012
		(0.045)		(0.046)		(0.051)		(0.056)		(0.056)
Female X Hispanic		-0.051		0.079		0.007		-0.053		0.003
		(0.042)		(0.042)		(0.044)		(0.047)		(0.048)
Female X Asian American		-0.068		0.002		0.033		0.056		0.000
		(0.037)		(0.031)		(0.031)		(0.034)		(0.034)
Female x Missing		-0.066		0.010		0.006		0.062		0.098
		(0.063)		(0.053)		(0.054)		(0.057)		(0.059)
Disadv X African American		-0.120		0.106		0.122		-0.053		0.009
		(0.061)		(0.062)		(0.066)		(0.072)		(0.072)
Disadv X Hispanic		-0.262		0.076		0.093		-0.033		0.186
		(0.060)		(0.060)		(0.061)		(0.066)		(0.068)
Disadv X Asian American		-0.092		-0.073		0.017		0.002		0.126
		(0.061)		(0.057)		(0.057)		(0.061)		(0.062)
Disadv X Missing		-0.008		0.020		0.041		-0.075		-0.128
		(0.111)		(0.103)		(0.104)		(0.112)		(0.116)
Observations	142728	136208	142728	136208	136958	130733	115618	110195	134341	128288
Pseudo R Sq.	0.161	0.565	0.041	0.128	0.03	0.142	0.029	0.137	0.046	0.185
Major, Dockets, Waiver, Early	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Academics, Nbhd/School, Ratings	N	Y	N	Y	N	Y	N	Y	N	Y

Source: Tables B.6.1R and B.6.2R of Document 415-9. Data restricted to non-ALDC applicants from the Classes of 2014–2019.

Notes: Standard errors below each coefficient in parentheses. "Sparse" means a model with relatively few covariates (Model 1 in the tables in Document 415-9); "Preferred" means the preferred model (Model 5 in the tables in Document 415-9).

Table F7: Personal and Overall Ratings, Selected Coefficients

	Personal		Alumni Personal		Overall		Alumni Overall	
	Sparse	Preferred	Sparse	Preferred	Sparse	Preferred	Sparse	Preferred
African American	-0.108	0.682	-0.132	0.236	-0.821	1.503	-0.664	0.126
	(0.025)	(0.053)	(0.021)	(0.041)	(0.019)	(0.038)	(0.020)	(0.040)
Hispanic	-0.075	0.279	-0.111	0.062	-0.237	0.922	-0.358	0.001
	(0.023)	(0.044)	(0.019)	(0.034)	(0.016)	(0.030)	(0.019)	(0.033)
Asian American	-0.346	-0.398	-0.010	-0.181	0.160	-0.136	0.232	0.160
	(0.018)	(0.034)	(0.014)	(0.025)	(0.012)	(0.022)	(0.014)	(0.024)
Missing	-0.237	-0.347	0.019	-0.129	0.095	-0.086	0.187	0.165
	(0.029)	(0.056)	(0.023)	(0.041)	(0.020)	(0.036)	(0.023)	(0.040)
Female	0.170	0.161	0.177	0.240	-0.017	0.117	-0.027	-0.094
	(0.014)	(0.039)	(0.011)	(0.032)	(0.010)	(0.027)	(0.011)	(0.031)
Disadvantaged	0.754	0.553	0.172	-0.075	0.603	0.743	0.191	0.068
	(0.026)	(0.052)	(0.022)	(0.044)	(0.019)	(0.038)	(0.021)	(0.043)
Female X African American		-0.239		-0.066		-0.163		-0.085
		(0.057)		(0.045)		(0.040)		(0.044)
Female X Hispanic		-0.015		-0.021		-0.013		-0.014
		(0.051)		(0.041)		(0.035)		(0.040)
Female X Asian American		0.095		0.053		0.040		-0.062
		(0.040)		(0.031)		(0.026)		(0.030)
Female x Missing		0.118		0.034		0.011		-0.041
		(0.069)		(0.054)		(0.045)		(0.052)
Disadv X African American		-0.324		0.101		-0.684		-0.066
		(0.073)		(0.061)		(0.053)		(0.059)
Disadv X Hispanic		-0.048		0.174		-0.353		-0.077
		(0.070)		(0.060)		(0.051)		(0.058)
Disadv X Asian American		0.058		0.087		0.100		-0.060
		(0.067)		(0.056)		(0.048)		(0.054)
Disadv X Missing		0.068		0.078		-0.155		-0.071
		(0.123)		(0.101)		(0.088)		(0.098)
Observations	142728	136208	111524	108054	142701	136183	111524	108054
Pseudo R Sq.	0.06	0.289	0.012	0.341	0.059	0.331	0.035	0.375
Major, Dockets, Waiver, Early	Y	Y	Y	Y	Y	Y	Y	Y
Academics, Nbhd/School, Ratings	N	Y	N	Y	N	Y	N	Y

Source: Tables B.6.3R and B.6.4R of Document 415-9. Data restricted to non-ALDC applicants from the Classes of 2014–2019.

Notes: Standard errors below each coefficient in parentheses. "Sparse" means a model with relatively few covariates (Model 1 in the tables in Document 415-9); "Preferred" means the preferred model (Model 5 in the tables in Document 415-9).

Table F8: The Role of Observed and Unobserved Factors in Racial/Ethnic Differences in Component Scores

	Overall	Academic	Extra-curricular	Teacher 1	Teacher 2	Counselor	Alumni Personal	Alumni Overall	Personal
Average Index Z-score (relative to White)									
African American	-1.129	-1.237	-0.663	-0.759	-0.722	-0.849	-0.253	-0.637	-0.374
Hispanic	-0.712	-0.791	-0.427	-0.451	-0.415	-0.514	-0.191	-0.421	-0.268
Asian American	0.120	0.259	0.109	0.142	0.116	0.049	0.027	0.073	0.020
Coefficients (White is normalized to zero)									
African American	1.503	-0.006	-0.217	0.012	0.104	0.164	0.236	0.126	0.682
Hispanic	0.922	-0.112	-0.146	-0.023	0.024	0.017	0.062	0.001	0.279
Asian American	-0.136	0.136	0.171	-0.159	-0.203	-0.095	-0.181	0.160	-0.398

Source: Table B.6.11R of Document 415-9.

Notes: The average index Z-score is calculated by taking the variables in the preferred ratings models absent race and admissions cycle and multiplying them by their corresponding coefficients from the ratings models. Then, the mean for white applicants is subtracted and we divide by the standard deviation. Finally, we take the averages for each racial group (note that mechanically this is zero for whites). Coefficients refer to the base race coefficients in the ratings models.

Table F9: Predicted and Actual Asian American Admit Rates by Admission Index Decile

	Distribution of Latent Admissions Index			
Decile	Actual	Normal	Flexible	Log-normal
Bottom 5	0.04	0.05	0.04	0.77
6	0.32	0.34	0.33	1.28
7	0.77	0.79	0.76	1.7
8	2.03	2.10	2.03	2.6
9	7.01	6.97	7.01	5.56
10	41.68	41.69	41.68	41.87

Notes: Actual refers to Table 9.1 of Document 415-9. Decile refers to the decile of the Asian American admissions index. Normal and Log-normal refers to the distribution of the admissions index. Flexible uses a normal distribution as well as the following transformations of the normal distribution: the square, the square interacted with the value being above zero, and the exponential. We obtained by Method of Simulated Moments the weights on these transformations that match to the actual distribution.

Table F10: Mother's and father's occupations vary in non-credible ways

	Admissions Class					
	2014	2015	2016	2017	2018	2019
Mother's Occupations						
Other	1266	4703	4339	4280	5666	5958
Homemaker	3476	4292	3967	4042	4629	3847
Unemployed	1449	2350	2274	2360	10	9
Low Skill.	1097	37	18	12	24	20
Self-Employed	0	991	989	928	1076	1138
Father's Occupations						
Other	1593	4608	4268	4587	4941	5663
Homemaker	44	56	50	61	101	71
Unemployed	963	1493	1390	1300	5	8
Low Skill.	1098	42	33	34	15	27
Self-Employed	0	2134	2148	2108	2335	2432

Source: Data presented in Table 3.2N of Document 415-9.

Notes: Construction of occupation categories described in Document 419-143.

Table F11: Intended Career varies in non-credible ways

	Admissions Class					
	2014	2015	2016	2017	2018	2019
Academic	1,723	25	19	15	2,247	13
Arts	846	331	321	284	390	283
Business	2,189	2,385	2,486	2,556	1,918	2,906
Communications	695	741	634	528	229	491
Design	283	161	131	101	82	105
Government	1,604	1,785	1,695	1,683	1,610	1,617
Health	234	95	85	107	4,944	96
Law	2,093	1,963	1,787	1,639	708	1,484
Library	63	0	0	0	0	0
Medicine	6,254	6,185	5,879	5,863	3	5,977
Religion	42	2	0	1	0	0
Science	3,268	5,242	5,437	5,519	9,182	7,394
Trade	2	7	8	7	6	9
Social Service	339	41	51	47	0	52
Teaching	167	660	598	598	17	514
Other	445	1,275	1,210	1,223	231	1,857
Undecided	1,821	5,022	4,614	4,887	3,537	3,661
Unknown	121	87	82	55	102	101
Total	22,189	26,007	25,037	25,113	25,206	26,560

Source: Data presented in Table B.4.1N of Document 415-9.

Table F12: Is a Pooled Model that finds no Asian American Penalty Robust?

	Average Marginal Effect
(1) Baseline pooled model from Card	-0.22%
(2) Interact race and disadvantaged	-0.32%*
(3) White and Asian American applicants only	-0.34%*
(4) Non-disadvantaged applicants only	-0.35%*
(5) Exclude personal rating	-0.65%*
(6) Exclude parental occupation	-0.37%*
(7) Combine (2), (5), & (6)	-0.95%*

Source: Data presented in Table 4.1N of Document 415-9. *=statistically different from zero at the 95% level. Marginal effects are calculated without perfect predictions.

Notes: All models exclude ALDC applicants.

Table F13: Is a Yearly Model that finds no Asian American Penalty Robust?

	Yearly with Card extracurriculars	Yearly with corrected extracurriculars	Pooled
(1) Baseline model from Card	-0.18%	-0.24%	-0.22%
(2) Interact race and disadvantaged	-0.29%	-0.36%*	-0.32%*
(3) White and Asian American applicants only	-0.37%	-0.47%*	-0.34%*
(4) Non-disadvantaged applicants only	-0.29%	-0.37%*	-0.35%*
(5) Exclude personal rating	-0.56%*	-0.62%*	-0.65%*
(6) Exclude parental occupation	-0.39%*	-0.47% *	-0.37%*
(7) Combine (2), (5), & (6)	-0.90%*	-0.98%*	-0.95%*

Source: Data presented in Table 4.2N of Document 415-9. *=statistically different from zero at the 95% level. Marginal effects are calculated without perfect predictions.

Notes: All models exclude ALDC applicants.

Table F14: Yearly estimates of the Asian American Penalty

	(1)	(2)	(3)	(4)	(5)
	Card Baseline	Interact Disadvantaged	No Personal Rating	No Parental Occupation	(2), (3), and (4)
2014	-0.31%	-0.38%	-0.79%	-0.69%	-1.23%
2015	-0.33%	-0.41%	-0.74%	-0.60%	-1.07%
2016	-0.02%	-0.16%	-0.72%	-0.34%	-1.12%
2017	-0.23%	-0.30%	-0.34%	-0.32%	-0.64%
2018	-0.57%	-0.71%	-0.97%	-0.75%	-1.33%
2019	0.37%	0.22%	0.19%	0.34 %	-0.03%
Avg. without 2019	-0.29%	-0.39%*	-0.71%*	-0.54%*	-1.08%*

Source: Data presented in Table 4.3N of Document 415-9. *=statistically different from zero at the 95% level. Marginal effects are calculated without perfect predictions.

Notes: All models exclude ALDC applicants.