Asian American Discrimination in Harvard Admissions*

Peter Arcidiacono
Duke University
NBER & IZA

Josh Kinsler
University of Georgia

Tyler Ransom
University of Oklahoma
& IZA

April 20, 2020

Abstract

Detecting racial discrimination using observational data is challenging because of the presence of unobservables that may be correlated with race. Using data made public in the *SFFA v. Harvard* case, we estimate discrimination in a setting where this concern is mitigated. Namely, we show that there is a substantial penalty against Asian Americans in admissions with limited scope for omitted variables to overturn the result. This is 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. Our preferred model shows that Asian Americans would be admitted at a rate 19% higher absent this penalty. Controlling for one of the primary channels through which Asian American applicants are discriminated against—the personal rating—cuts the Asian American penalty by less than half, still leaving a substantial penalty.

*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

Discrimination on the basis of race or ethnicity has a long and sordid history in the United States, with the attacks against Asian Americans during the COVID-19 pandemic providing a recent example (Tavernise and Oppel Jr., 2020). While significant progress has been made, unequal treatment on the basis of race or ethnicity—be it in the labor market, criminal justice, or the education system—remains a concern.[1] But detecting how discrimination affects these and other outcomes using observational data is difficult because of the presence of unobservables that may be correlated with race.

Concern regarding unobserved attributes is particularly salient in this context for two reasons. First, adding control variables to a model that links race to a specific outcome often reduces racial disparities.[2] This occurs in part because race can proxy for differences in unobserved factors, such as pre-labor-market or pre-college human capital. If additional measures of these underlying attributes were available and included in the model, the evidence for discrimination might be eliminated. Second, unease about unobserved attributes is magnified when a statistical model explains only a small portion of the outcome. Most observational studies of discrimination explain significantly less than half of the variation in the relevant outcome, leaving substantial room for omitted variable bias. Concern about omitted variable bias has resulted in a shift to audit and correspondence studies that rely on fictitious applicants.[3] However, these studies also face limitations such as experimenter effects, heterogeneity in unobservables, where they can be applied, and how applicants signal race.[4]

[1]A lengthy literature in economics examines racial discrimination in the workplace (e.g. Bertrand and Mullainathan, 2004; Charles and Guryan, 2008), criminal justice system (e.g. Knowles, Persico, and Todd, 2001; Anwar and Fang, 2006; Antonovics and Knight, 2009), and education system (e.g. Dorsey and Colliver, 1995; Figlio, 2005; Botelho, Madeira, and Rangel, 2015).

[2]See, for example, Altonji and Blank (1999) and Arcidiacono, Aucejo, and Spenner (2012), where adding controls reduces racial disparities in earnings and college grades, respectively.

[3]Examples include Bertrand and Mullainathan (2004), Edelman, Luca, and Svirsky (2017), and Agan and Starr (2018).

[4]See Heckman (1998) and Neumark (2018) for more detail on experimenter effects and heterogeneous variances in the unobserved components of productivity. Typically audit and correspondence studies must be implemented in environments where some anonymity can be preserved (so it will not be clear that it is an experiment) and outcomes that are early in the process (for example, callbacks not wages). Finally, race is often communicated through an applicant's name which complicates interpretation since distinctively African American names can also indicate socioeconomic status (Fryer and Levitt, 2004). An exception to this criticism is Pope and Sydnor (2011), who rely on pictures to indicate applicant race.

Using data made public in the *SFFA v. Harvard* case, we study discrimination using observational data in an environment where these two concerns are mitigated: discrimination against Asian Americans in college admissions. Namely, we show that there is a substantial penalty against Asian Americans in admissions with very limited scope for omitted variables to overturn the result. This is 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.

For years there has been a perception of discrimination against Asian Americans in elite college admissions (Golden, 2006; Fuchs, 2019). But 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 examine how Asian American applicants are treated relative to their white counterparts.

The data we analyze covers the admissions cycles for applicants who, if they were to have graduated in four years, would have done so as the Classes of 2014–2019. In addition to many demographic and academic measures, the data include information on the various Harvard ratings that influence admissions decisions. These include Harvard admissions officers' ratings of the applicant 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, it includes information on alumni interviews of the applicants in the form of an overall score and a personal score. We describe the data and the admissions process itself in Section 2.

There are certain groups of applicants that receive substantial preferences in admissions. These include recruited athletes, legacies, children of donors, and children of faculty and staff (ALDC).[5] We focus on applicants that are not in one of these special recruiting categories so that we can compare similarly situated applicants. Typical applicants (i.e., those who are not ALDC) make up over 95% of domestic applicants and over 97% of Asian American

[5] See Section 2.1.3 of Arcidiacono, Kinsler, and Ransom (2019b) for how ALDC applicants are treated differently in Harvard's admissions process.

applicants. It should be clear that consistent estimates of racial discrimination against typical Asian American applicants can be obtained from estimating an admissions model only on typical applicants. However, given the central role ALDC applicants played in the court case, we show both theoretically and empirically why estimating without this group is appropriate in Section 3.[6] For ease of exposition, when we refer to Asian American applicants going forward, we mean the subset who are not ALDC.

After removing ALDC applicants, Asian Americans actually have a slightly higher unconditional admit rate than whites. But as we show in Section 4, 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%.

Of course, Harvard values much more than just academics. And here, too, Asian American applicants as a whole perform as well or better than white applicants on almost all of Harvard's ratings.[7] The strong performance of Asian Americans is in part because academically strong applicants also tend to be stronger on Harvard's ratings.[8]

But Harvard's ratings may also be affected by racial preferences and penalties. Indeed, Harvard acknowledges that race is one of the inputs into the overall rating. Consistent with this, we find substantial preferences given to African Americans and Hispanics and a corresponding penalty for Asian Americans in the overall rating. But we also find that race substantially influences the personal rating, to the detriment of Asian Americans.

The arguments for the personal rating being influenced by race—with Asian Americans receiving a penalty—are abundant. We list three here, though we also provide additional evidence in Section 5. First, Asian American applicants are slightly stronger than white applicants on the observable characteristics correlated with the personal rating, yet receive

[6] Note that estimating only on typical applicants is equivalent to estimating a model where ALDC status is fully interacted with the controls when the purpose is to measure the effect of race for typical applicants.

[7] The two exceptions are the athletic and personal ratings.

[8] This is true for all ratings but one: there is no information in the public record regarding how the athletic rating—at least for non-recruited athletes—is correlated with academic strength.

markedly lower personal ratings. Second, alumni interviewers—who are primarily white but actually met the applicants in person—score Asian American applicants better on the personal rating than Harvard admissions officers do.[9] Finally, there is clear evidence that other racial groups receive a tip on the personal rating. For example, conditional on observables, African Americans receive substantially higher personal ratings despite having observable characteristics associated with lower personal ratings. Indeed, African Americans in the top ten percent of the applicant pool according to grades and test scores are over twice as likely as their Asian American counterparts to receive a strong personal rating.

Given our findings that race influences both the overall and personal ratings, our preferred model of admissions, described in Section 6, 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 past work on admissions and, as a result, the model fits the data extremely well.[10]

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%.[11] Yet, even if the personal rating is included in the model, the statistical case for discrimination against Asian Americans is still clear and the penalty remains large.[12]

Concerns about the scope for omitted variable bias are significantly reduced in our setting. The richness of the applicant data yields a model of admissions outcomes that matches

[9] Less than 1% of typical Asian American applicants receive an interview with an admissions officer. See Table 7.3N of Document 415-9.

[10] Previous work on college admissions typically uses data from a third party, and thus does not have access to detailed admissions data. As a result, the admissions models employed often only include basic measures 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.

[11] 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).

[12] The average marginal effect is -0.54 percentage points, a little more than half the penalty in the preferred model. Indeed, as we show in Appendix B, the penalty Asian Americans receive is robust to a host of alternative modeling assumptions.

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 and preferences for legacies and athletes in elite college admissions and obtain a Pseudo R^2 of around 0.2. In a setting analogous to ours—elite academic journal editorial decisions—Card et al. (2020) and Card and DellaVigna (2020) estimate rich models of revise and resubmit decisions for a sample of manuscripts. Their models are notable in that they incorporate reviewer recommendations—which we would expect to be highly correlated with editor decisions—yet the Pseudo R^2 of their models is never larger than 0.5. In Section 6 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. First, Asian American applicants are significantly stronger than white applicants on the observable characteristics—outside of race—that Harvard values when making admissions decisions.[13] 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. 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). However, in our case, the estimated effect of being Asian American is negative *despite* Asian Americans being positively selected on the observables associated with admission. Thus, if we follow the literature and assume Asian American applicants are also stronger on the unobserved factors affecting admission, the actual degree of discrimination is larger than our estimates indicate.

A second reason that we are likely understating the penalty against Asian American applicants is that some of the other Harvard ratings we include also show patterns indicative of discrimination against Asian American applicants.[14] Similar to the personal rating, how

[13] They are also at least as strong or stronger than white applicants on the non-academic portion of the observables. See Section 6.

[14] A common issue researchers face is whether to control for variables that are themselves possibly a result

Harvard admissions officers rate the high school teacher and counselor letters also suggests racial bias. Namely, Asian Americans are stronger on the observable characteristics associated with these ratings, yet Asian Americans receive lower scores on these ratings conditional on the observables. We include these ratings in our preferred model as the penalties are not as egregious as what is observed in the personal rating. In doing so, we set an extremely high bar for establishing discrimination, likely underestimating the penalty Asian American applicants face as a result.

That Asian Americans still manage to make up over 20% of Harvard admits despite the disadvantages they face in the admissions process is striking. But it does not justify Asian Americans losing out on a substantial number of seats because of their race and their unexplainably low personal ratings.

2 Data and Admissions Process

We now briefly describe the data used in our analysis, as well as how the admissions process operates at Harvard. More details are available in Section 2 of Arcidiacono, Kinsler, and Ransom (2019b).

2.1 Data

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).[15]

In all, the data used in Document 415-9 consist of 142,728 typical domestic applicants who have complete application data. For each applicant, the microdata contain details

of racial bias. Neal and Johnson (1996) pointed this out when studying racial differences in the labor market and researchers have addressed similar issues in racial discrimination in criminal justice and school discipline (see Rehavi and Starr (2014) and Kinsler (2011), respectively).

[15] For a complete list of legal documents we rely on in this paper, see Appendix Table D1.

on academics, demographics, ALDC status, Harvard internal ratings, and a host of other variables (e.g. whether the applicant applied for financial aid, parental education, etc.).

2.2 Admissions Process

In each admissions cycle, applications are divided into geographical areas called dockets. Harvard admissions officers are divided into subcommittees according to each docket. Admissions officers read each application and assign three classes of ratings on a number of dimensions: (*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. Ratings typically take on values between 1 and 5, with 1 being the best. For some ratings, +/- suffixes are used to distinguish values within a given integer. For example, a 1- is a better rating than a 2+.

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. For the athletic rating, high ratings appear to be related to one's ability to play competitively for Harvard's teams, either as a recruit or a walk-on. The personal rating, which is meant to capture personal qualities such as likeability, courage, and kindness, is the most vague.[16] 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."[17]

It is also important to note that the school support ratings are assigned by the Harvard admissions office and not directly by the high school teachers themselves. High school

[16] 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.

[17] Harvard's admissions director said the following when asked about the personal rating criteria: "[The guidelines are] not terribly helpful. I think 'unusual appeal' is ... the phrase we use [for a 1]. ... [R]eaders will construe that in different ways. And 2 is ... a very attractive person to be with and have in your school community and widely respected. And 3 means, you know, just fine. ... 4 is a kind of negative" (Document 419-1, p. 171).

teachers submit a letter of recommendation and also fill out a questionnaire where they are asked to rate the applicant's academic strength and personal qualities (such as maturity, motivation, integrity and leadership skills).[18] Harvard admissions officers then assign the school support ratings based on the information contained in recommendation letters and questionnaire.

In addition to submitting application materials, 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. A small number of applicants interview with Harvard admissions staff.

After applications have been read and assigned ratings for each category, a subset of the applicants are passed on to an additional reader, called the "third reader."[19] Provisional admissions decisions are then made at the docket subcommittee level. In March, final decisions are made at the full committee level with all admissions officers present.[20]

3 Modeling Admissions and the Relevant Sample

As noted in the introduction, we focus on estimating the Asian American penalty among typical applicants. 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.[21] As we show, these assumptions are violated in the Harvard data.

[18]For an example of this taken from the Class of 2012 Casebook, see Trial Exhibit DX 002, pp. 6–7.

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

[20]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`.

[21]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 has no effect on the estimates of racial discrimination against typical applicants.

Discrimination can manifest itself in many forms and it may operate in some places more than in others. Consider professional football players. Well after the National Football League had integrated, there was still substantial discrimination against black quarterbacks, with the perception being that they were not cerebral enough or did not have the leadership skills necessary for the position (Hruby, 2019). Clearly it would be insufficient for a team to refute an accusation of discrimination against quarterbacks by showing it does not discriminate against non-quarterbacks.

A similar dynamic can be at work between ALDC and typical applicants. Asian Americans are stereotyped as enjoying math, aspiring to be doctors, and being a "model minority" (Kim and Lewis, 1994; Yam, 2019). One way of overcoming those stereotypes is through connections, such as having the opportunity to make oneself known to the admissions committee or Harvard in general. Athletes provide the clearest example through their contact with Harvard coaches. Chances to interact with the Harvard community will also be present for legacies through alumni events. More importantly, simply being in one of these ALDC categories bucks the stereotype given how few Asian Americans are ALDC: less than 2.5% of Asian American applicants fall into one of these groups, compared to over 9% of whites (see Table 3.2 of Document 415-8).

Yet, regardless of whether *ALDC* Asian Americans are discriminated against relative to their white counterparts, whether Harvard is discriminating against *typical* Asian Americans can be determined from looking only at typical applicants. To illustrate, 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^*.[22] 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

[22] 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.

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

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

(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.[23] 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)

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

ALDC, we can specify the index as:

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

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

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

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 (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 we will show, the evidence suggests that the effects of being Asian American differ for ALDC applicants *and* the effects of X_i differ for ALDC applicants. One example is academics. Preliminary evidence that academics matter differently depending on ALDC status can be seen in Tables 2 and 6 of Arcidiacono, Kinsler, and Ransom (2019b).[24] The former shows that those who receive an academic rating of 5 are guaranteed rejection unless they are athletes, and those who receive a 4 are virtually guaranteed rejection unless they are ALDC. The latter shows that, in the bottom decile of the academic index (a weighted average of SAT scores and high school GPA), 6 percent of white LDC applicants were admitted, yet *zero* typical white applicants were admitted from this decile.

We will also show that Asian Americans are substantially stronger than white applicants on academics. If including ALDC applicants results in the coefficients on academics being

[24] For additional evidence of race mattering differently for athletes and legacies that generalizes beyond Harvard, see Table 6 of Espenshade, Chung, and Walling (2004).

attenuated (assuming no interactions between ALDC status and academics), this will make it appear as though academics are not as important as they really are for typical applicants. Hence, Harvard could falsely claim they were not discriminating against Asian Americans; academics are just not as important as one would think despite the important role academics play for typical applicants.

Given the substantial difference in admit rates depending on ALDC status as well as the evidence that academics matter less for this group, we focus on typical applicants.[25] We return to show how including ALDC applicants distorts the coefficients of the model in Section 6.2.

4 Descriptive Analysis

We now turn to the characteristics of typical Asian American and white applicants. 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.

4.1 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.[26] Those who are labeled disadvantaged

[25] As further evidence that admissions for ALDCs operates differently, Arcidiacono, Kinsler, and Ransom (2019a) show that the share of admits who are legacies or recruited athletes (LA) shows no time trend for the Classes of 2000–2017, despite the share of applicants who are LA showing a steep downward time trend over the same period (see Figure 2).

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

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%).[27] Asian Americans are also more 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.

With regard to parental education, Asian American mothers are slightly less likely (by 0.3 percentage points) to have a degree higher than a bachelor's, while Asian American fathers are more likely (by 8.5 percentage points) to have a degree higher than a bachelor's. Both of these variables are positively correlated with admission.

4.2 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 SAT math and SAT II 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. 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.[28] The only measure of academic preparation that whites do comparably

[27]These numbers are computed using Bayes' rule from figures reported in Table 1. For example:

$$\begin{aligned}\Pr(\text{Admit} \mid \text{Disadvantaged}) &= \frac{\Pr(\text{Disadvantaged} \mid \text{Admit})\Pr(\text{Admit})}{\Pr(\text{Disadvantaged})}\\ &= \frac{0.2186 \times 0.0513}{0.1085}\\ &= 10.33\%\end{aligned}$$

for Asian Americans. A similar strategy yields the admit rate for disadvantaged whites and non-disadvantaged applicants.

[28]The academic index is a weighted average of the applicant's scores on the SAT, SAT II, and high school grade point average (or class rank). It is used by Ivy League institutions to ensure recruited athletes meet minimum academic standards. See Document 415-8 footnote 29 for a more detailed discussion of the academic index.

on is the SAT verbal.

The differences in academic achievement between white and Asian American applicants become even more staggering when looking at deciles of the academic index. The first set of columns of 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.[29] 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.

4.3 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

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

admitted would increase by 828, a 40% increase.[30] 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.

In Table 3, we report how white and Asian American applicants score on each of Harvard's ratings. 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 2 or better, 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 a 2 or better 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 a 2 or better 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.[31] Receiving an athletic rating of 2 does boost one's chances of admission. However, the change in admissions probability from moving from a 3 to a 2 is much smaller than the corresponding change in any of the other ratings, and the share of applicants who receive a 2 on the athletic rating is also smaller than the corresponding share on any other rating. And there is no difference in the admit rate between those who get a 3 on the athletic rating versus those who get a 4, while for the other three profile ratings, getting a four virtually guarantees rejection. Thus, it appears that the athletic rating is less important than the other profile ratings in determining admissions outcomes.

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.[32] 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.[33]

[30] This number is obtained by fixing the total number of white and Asian American admits at their current level, 4884, and then randomly sampling from the tenth decile to fill the class.

[31] When the Office of Civil Rights investigated Harvard, it was these two ratings that were found to be the most subjective. See Trial Exhibit P555.

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

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

4.4 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.[34] Appendix Tables D2 through D4 show the share of applicants receiving a 2 or better for each of the ratings by academic 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 declines significantly from the 10th to the 1st 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 shares and then 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.[35] 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, there 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

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

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

relative to the three other major racial/ethnic groups.[36] The first column of Table 4 shows the share of Asian Americans in the top academic index decile who receive a 2 or better 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 2 or better on the personal rating was over seven percentage points higher than Asian Americans, a 33% increase. While whites in the tenth decile scored higher on other ratings as well, the gaps are smaller.

The fourth column shows the lowest decile in the white academic index distribution that would still have a higher probability of receiving a 2 or better rating than Asian Americans. For the personal rating, whites in the 6th academic index decile have a higher probability of receiving a 2 or better 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 2 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 2 or better 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. African American applicants in the top (10th) decile are also twice as likely to receive a 2 or better as African American applicants in the third decile.[37] Yet, even African Americans in the *third* academic index decile score better 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

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

[37] See Appendix Table D3.

decile. Like the personal rating, Asian Americans are rated the lowest and this is 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. In the next section, we estimate models of these ratings to see the effects of race after accounting for many of the multitude of variables available in the Harvard database.

5 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 formally 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.[38] 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.[39]

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

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

The ordered logit framework is appropriate when the dependent variable, in this case an applicant rating, takes on a discrete number of values that have a natural ordering. In Harvard's rating system, applicants are typically rated on a scale from 1 to 4, with 1 being the best rating.[40] Suppose there is a latent index by which admissions officers rate applicants, where R indexes the rating (e.g. $R \in \{\text{Overall}, \text{Academic}, \text{Personal}, \ldots\}$):

$$\pi_i^R = X_i^R \gamma^R + \varepsilon_i^R \tag{4}$$

such that the observed rating for applicant i, Y_i^R, takes on a particular value, say 3, when π_i^R is within a certain range. Namely:

$$Y_i^R = \begin{cases} 1 & \text{if } \pi_i^R \geq k_1^R \\ 2 & \text{if } k_1^R > \pi_i^R \geq k_2^R \\ 3 & \text{if } k_2^R > \pi_i^R \geq k_3^R \\ 4 & \text{if } k_3^R > \pi_i^R \end{cases} \tag{5}$$

where $k_1^R > k_2^R > k_3^R$ are the thresholds associated with each ranking. Assuming that ε_i^R follows a Type 1 extreme value distribution, this leads to an ordered logit model. The probabilities of receiving each of the ratings conditional on X_i are given by:

$$\begin{aligned}
\Pr\left(Y_i^R = 4\right) &= \frac{\exp\left(k_3 - X_i^R \gamma^R\right)}{1 + \exp\left(k_3 - X_i^R \gamma^R\right)} \\
\Pr\left(Y_i^R = 3\right) &= \frac{\exp\left(k_2 - X_i^R \gamma^R\right)}{1 + \exp\left(k_2 - X_i^R \gamma^R\right)} - \frac{\exp\left(k_3 - X_i^R \gamma^R\right)}{1 + \exp\left(k_3 - X_i^R \gamma^R\right)} \\
\Pr\left(Y_i^R = 2\right) &= \frac{\exp\left(k_1 - X_i^R \gamma^R\right)}{1 + \exp\left(k_1 - X_i^R \gamma^R\right)} - \frac{\exp\left(k_2 - X_i^R \gamma^R\right)}{1 + \exp\left(k_2 - X_i^R \gamma^R\right)} \\
\Pr\left(Y_i^R = 1\right) &= 1 - \frac{\exp\left(k_1 - X_i^R \gamma^R\right)}{1 + \exp\left(k_1 - X_i^R \gamma^R\right)}
\end{aligned} \tag{6}$$

To test whether discrimination is present in the ratings, we examine whether the γ^R's corresponding to each of the race dummies are statistically different from zero (where whites serve as the reference group). Selected coefficients from the ratings models are presented

415-9.

[40] The extracurricular rating includes a rating of 5 which indicates a significant family or work responsibility. Approximately 0.5% of applicants obtain this rating. For ease of exposition, we describe the ratings models as if they all run on a 4-point scale. When we estimate the extracurricular rating model, we incorporate this fifth category.

in Appendix Tables D5 and D6. 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.[41] 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 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 in the full model than in the sparse model.[42] 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 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%).[43]

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. Figure 2 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

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

[42]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.

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

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 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 according to $X_i^R\hat{\gamma}^R - X_{i,race}^R\hat{\gamma}_{race}^R - X_{i,yr}^R\hat{\gamma}_{yr}^R$. We then subtract off the white mean 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 D7.[44] In Figure 3 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 related to the number of AP exams, AP exam scores, and academic awards.[45] We know from Table 1 that Asian Americans are stronger on AP exams, and are likely stronger on other unobserved academic measures.[46]

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.[47]

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.

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

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

[46] 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.

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

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 Figure 3, there is also evidence that Asian American applicants are discriminated against in the school support ratings and alumni personal rating. 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 see muted preferences for African American applicants who were female or disadvantaged.[48] The only other rating that has this pattern is the overall rating, a rating that we know Harvard uses to directly implement preferences.

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

6 Discrimination in Harvard Admissions

The descriptive evidence related to admissions in conjunction with the ratings analysis suggests that there is scope for discrimination against Asian Americans 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, discuss the estimated Asian American penalty, and analyze the potential for omitted variable bias.

6.1 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.[49] 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.[50] 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

[49] A similar discussion of our modeling approach is presented in Arcidiacono, Kinsler, and Ransom (2019b), Section 4 and Appendix C.

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

generate similarly constituted admit classes over time, even if the applicant pool is changing. 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.

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 is also influenced by racial preferences.

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.[51] 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 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

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

are excluded. The full set of controls for our preferred model is listed in Appendix A.

6.2 Estimates of the Asian American Admissions Penalty

A subset of the estimated parameters for our preferred admissions model is displayed in Table 5.[52] The coefficient on the Asian American indicator is negative and significant (-0.466), indicating that, all else equal, Asian American applicants are penalized relative to white applicants. While the estimated Asian American penalty does not vary significantly with either disadvantaged or early applicant status, 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 underrepresented minority groups and disadvantaged students. Applicants who receive a two or better on the academic, extracurricular, or personal ratings are significantly more likely to be accepted. Applicants who receive a four on any of these criteria see their chances of admission diminished relative to receiving a three, though the effect on the athletic rating is an order of magnitude smaller.[53] The coefficients associated with being African American, Hispanic, or disadvantaged are all large, positive, and statistically significant.[54] Finally, conditional on all other observed attributes, early action applicants are more likely to be admitted than their regular admissions counterparts.

The parameter estimates listed in Table 5 represent how Harvard values various charac-

[52]Table D8 illustrates how the race-related parameters change as we build up the set of included applicant attributes. While it is not possible to directly compare the magnitude of the coefficients across models (since the variance of the unexplained portion of the admissions decision is shrinking and this tends to inflate the estimated parameters) there are interesting patterns. For example, when we control only for race and a handful of other demographic characteristics, the coefficient on the Asian American indicator is positive 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.

[53]Obtaining a rating of five on the extracurricular and athletic 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).

[54]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.

teristics among typical applicants. As discussed at length in Section 3, we exclude ALDC applicants from the bulk of our analysis since there is both qualitative and quantitative evidence that the admissions process works differently for this special set. For example, recruited athletes have an advocate at the university outside of the admissions office, and in the past the Dean of Admissions would review every legacy applicant—benefits typical applicants do not receive.[55] Harvard also appears to value applicant attributes differently according to ALDC status. As evidence of this, Table 6 shows how the coefficients associated with Harvard's academic and extracurricular rating change as LDC and recruited athlete applicants are added to the estimation sample.[56] The most striking change relative to a model with only typical applicants is the penalty an applicant receives for obtaining a 4 rating for either the academic or extracurricular rating. When ALDC applicants are included, the penalties for receiving a 4 are much smaller, indicating that the academic and extracurricular ratings are less important for these special groups. The broader point is that if Harvard valued the attributes of typical and ALDC applicants similarly, there would have been no change in the coefficients. The fact that the coefficients change so significantly indicates that the appropriate way to estimate how Harvard values attributes among typical applicants is to either interact ALDC status with all applicant attributes, or more simply, exclude ALDC applicants.[57]

The parameter estimates from our preferred specification that includes only typical applicants indicate the existence of an Asian American penalty. However, 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 can show how the probability of admission would change for Asian

[55] See page 47 of Document 419-143 for an example of a coach advocating for her players with the admissions office. See Golden (2006, p. 27) for a statement by Harvard's Dean of Admissions that he personally reviews every legacy's application. For a discussion of the benefits universities receive from alumni and athletic connections, see Meer and Rosen (2009a,b, 2010, 2012).

[56] The admissions model that includes ALDC applicants is identical to our preferred model, except that it includes separate indicators for recruited athlete, legacy, double legacy, child of faculty/staff, and dean's interest list. These indicators are also interacted with race and admissions cycle. Coefficients are presented in Table B.7.2R of Document 415-9 and Table 2 of Exhibit 287.

[57] Note that even when ALDC applicants are added to estimation sample in an inappropriate fashion by only incorporating ALDC indicators and interacting ALDC with race, the average marginal effect for being Asian American across all applicants is still negative and significant at -0.80 (see Exhibit 287, paragraph 9 on p. 4).

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 X. The index of observables, Z, for this applicant according to the log odds formula is given by

$$Z = \ln\left(\frac{X}{1-X}\right) \tag{7}$$

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 7 lists the the results of these transformation exercises. The first entry in the 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.[58] 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%.[59] In Table 8 we report the average marginal effect of being an

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

[59]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

Asian American applicant by admissions index decile.[60] 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%.

6.3 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 the 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 5, we presented evidence that the personal rating is significantly influenced by applicant race, and is thus also an improper control.

We now include the personal rating in the model to see how much of the Asian American penalty operates through it. Adding the personal rating still 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. 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.[61] A more reasonable interpretation is that the reduction in the Asian American penalty shows how much of the penalty operates through the personal rating, implying that the penalties through the personal rating account for a

-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.

[60] 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.

[61] The Asian American penalty is robust to a number of other modeling decisions. As we show in Appendix B, the penalty only becomes insignificant when one make a series of questionable modeling choices, all of which entail assuming there is no bias against Asian Americans in the personal rating.

little less than half of the total Asian American penalty.

6.4 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 interviews. Relative to previously published analyses of admissions decisions, the richness of the available data is without rival.[62] 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. The evidence presented in the following sections suggests this is unlikely.

6.4.1 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, similar to our analysis of racial preferences in Harvard's ratings, 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 the admissions index, with higher deciles associated with stronger observed characteristics. The results for Asian American and white applicants are displayed in the first two columns of Table 9. Asian American applicants are stronger, with 13.1% of Asian American applicants in the top decile and only 10.5% of

[62] Various papers have explored the impact of race (Long, 2004) and legacy status (Espenshade and Chung, 2005; Hurwitz, 2011) on 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 U.K. university. This paper is closest to ours in the richness of the available data. However, the authors do not report the fit of their admissions models.

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%.[63]

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.[64] However, Asian American applicants could be weaker along observed non-academic 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.[65] 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 9. 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.[66] 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

[63] 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 8 imply that over 93% of Asian American admissions come from those in the top 20% of the observed admissions index.

[64] Table 5.3 of Document 415-8 shows that over 51% of the class would be Asian American if admissions were based entirely on the academic index.

[65] 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.

[66] Document 415-9 section 9.3 shows that dockets 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.

Asian American applicants are just as strong, if not stronger than white applicants.

6.4.2 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 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.[67] 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 $AI_i = \alpha A_i + X_i\beta$, implying:

$$Y_i^* = AI_i + \epsilon_i. \tag{8}$$

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 (8), 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.[68] Assuming that the distribution of AI_i

[67] 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.

[68] This distribution could be calculated using the underlying data. However, the information needed to do

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 C, 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 show in Appendix C, 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. In comparison, an admissions model with no controls would produce an accuracy for admits equal to 0.3%.[69]

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.

7 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.[70]

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

so is not in the public record.

[69] An admissions model with no controls would randomly assign 5.45% of applicants as admits and 94.55% as rejects. The accuracy rate for admits would then be given by $5.45\% \times 5.45\% = 0.3\%$.

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

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.[71] 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.

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.

[71] 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.

References

Agan, Amanda and Sonja Starr. 2018. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *Quarterly Journal of Economics* 133 (1):191–235. 2

Altonji, Joseph G. and Rebecca M. Blank. 1999. "Race and Gender in the Labor Market." In *Handbook of Labor Economics*, vol. 3, chap. 48. Elsevier, 3143–3259. 2

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. 6

Antonovics, Kate and Brian G. Knight. 2009. "A New Look at Racial Profiling: Evidence from the Boston Police Department." *Review of Economics and Statistics* 91 (1):163–177. 2

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. 2

Arcidiacono, Peter. 2005. "Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica* 73 (5):1477–1524. 26

Arcidiacono, Peter, Esteban M. Aucejo, and Ken Spenner. 2012. "What Happens after Enrollment? An Analysis of the Time Path of Racial Differences in GPA and Major Choice." *IZA Journal of Labor Economics* 1 (1):5. 2

Arcidiacono, Peter, Josh Kinsler, and Tyler Ransom. 2019a. "Divergent: The Time Path of Legacy and Athlete Admissions at Harvard." Working paper, Duke University. 13

———. 2019b. "Legacy and Athlete Preferences at Harvard." Working paper, Duke University. 3, 7, 12, 25, 49, 56

———. 2019c. "Recruit to Reject? Harvard and African American Applicants." Working paper, Duke University. 64, 65

Bertrand, Marianne and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94 (4):991–1013. 2

Bhattacharya, Debopam, Shin Kanaya, and Margaret Stevens. 2017. "Are University Admissions Academically Fair?" *Review of Economics and Statistics* 99 (3):449–464. 5, 31

Blau, Francine D. and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3):789–865. 5

Botelho, Fernando, Ricardo A. Madeira, and Marcos A. Rangel. 2015. "Racial Discrimination in Grading: Evidence from Brazil." *American Economic Journal: Applied Economics* 7 (4):37–52. 2

Card, David and Stefano DellaVigna. 2020. "What Do Editors Maximize? Evidence from Four Economics Journals." *Review of Economics and Statistics* 102 (1):195–217. 6

Card, David, Stefano DellaVigna, Patricia Funk, and Nagore Iriberri. 2020. "Are Referees and Editors in Economics Gender Neutral?" *Quarterly Journal of Economics* 135 (1):269–327. 6

Charles, Kerwin Kofi and Jonathan Guryan. 2008. "Prejudice and Wages: An Empirical Assessment of Becker's *The Economics of Discrimination*." *Journal of Political Economy* 116 (6):773–809. 2

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. 10, 14, 19, 32, 51, 63

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. 5, 7, 16, 19, 21, 23, 24, 26, 28, 29, 32, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 54, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73

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. 8, 63

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. 54, 55, 56, 57, 58, 63, 72, 73

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. 28, 54, 55, 63, 71

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. 17, 23, 54, 63

Dorsey, J. Kevin and Jerry A. Colliver. 1995. "Effect of Anonymous Test Grading on Passing Rates as Related to Gender and Race." *Academic Medicine* 70 (4):321–323. 2

Edelman, Benjamin, Michael Luca, and Dan Svirsky. 2017. "Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment." *American Economic Journal: Applied Economics* 9 (2):1–22. 2

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. 34

Espenshade, Thomas J. and Chang Y. Chung. 2005. "The Opportunity Cost of Admission Preferences at Elite Universities." *Social Science Quarterly* 86 (2):293–305. 5, 31

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. 5, 6, 12

Exhibit 287. 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://www.courtlistener.com/recap/gov.uscourts.mad.165519/gov.uscourts.mad.165519.512.2_4.pdf`. Declaration outlining use of Bayes' rule to compute admissions probabilities of previous admits. 28, 49, 63

Figlio, David N. 2005. "Names, Expectations and the Black-White Test Score Gap." Working Paper 11195, National Bureau of Economic Research. 2

Fryer, Roland G. and Steven D. Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *Quarterly Journal of Economics* 119 (3):767–805. 2

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. 3

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. 3, 28

Heckman, James J. 1998. "Detecting Discrimination." *Journal of Economic Perspectives* 12 (2):101–116. 2

Hruby, Patrick. 2019. "'We've Come So Far': How Black QBs Defied a Racist Past to Become Football's Future." *The Guardian* URL `https://www.theguardian.com/sport/2019/sep/20/black-quarterbacks-history-stereotypes`. Accessed November 18, 2019. 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. 5, 31

Kim, Pan Suk and Gregory B. Lewis. 1994. "Asian Americans in the Public Service: Success, Diversity, and Discrimination." *Public Administration Review* 54 (3):285–290. URL http://www.jstor.org/stable/976733. 10

Kinsler, Josh. 2011. "Understanding the Black-White School Discipline Gap." *Economics of Education Review* 30 (6):1370–1383. 7

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. 2

Krauth, Brian. 2016. "Bounding a Linear Causal Effect Using Relative Correlation Restrictions." *Journal of Econometric Methods* 5 (1):117–141. 6

Long, Mark C. 2004. "Race and College Admissions: An Alternative to Affirmative Action?" *Review of Economics and Statistics* 86 (4):1020–1033. 5, 31

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. 6, 33

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. 33

Meer, Jonathan and Harvey S. Rosen. 2009a. "Altruism and the Child Cycle of Alumni Donations." *American Economic Journal: Economic Policy* 1 (1):258–286. 28

———. 2009b. "The Impact of Athletic Performance on Alumni Giving: An Analysis of Microdata." *Economics of Education Review* 28 (3):287–294. 28

———. 2010. "Family Bonding with Universities." *Research in Higher Education* 51 (7):641–658. 28

———. 2012. "Does Generosity Beget Generosity? Alumni Giving and Undergraduate Financial Aid." *Economics of Education Review* 31 (6):890–907. 28

Neal, Derek A. and William R. Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104 (5):869–895. 7

Neumark, David. 2018. "Experimental Research on Labor Market Discrimination." *Journal of Economic Literature* 56 (3):799–866. 2

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37 (2):187–204. 6

Pope, Devin G. and Justin R. Sydnor. 2011. "What's in a Picture? Evidence of Discrimination from Prosper.com." *Journal of Human Resources* 46 (1):53–92. 2

Rehavi, M. Marit and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6):1320–1354. 7

Tavernise, Sabrina and Richard A. Oppel Jr. 2020. "Spit On, Yelled At, Attacked: Chinese-Americans Fear for Their Safety." *The New York Times* URL https://www.nytimes.com/2020/03/23/us/chinese-coronavirus-racist-attacks.html. Accessed May 26, 2020. 2

Trial Exhibit DX 002. 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/D002.pdf. 2012 Casebook. 9, 63

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. 53, 54, 63

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. 53, 63

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. 23, 63

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. 8, 13, 27, 53, 54, 63

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. 53, 54, 63

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. 53, 63

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. 16, 63

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. 43, 63

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. 8, 53, 54, 63

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. 33

Yam, Kimmy. 2019. "Is Andrew Yang 'Reclaiming' Stereotypes with Asian Jokes? Experts Say Not So Much." *NBC News* URL https://www.nbcnews.com/news/asian-america/andrew-yang-reclaiming-stereotypes-asian-jokes-experts-say-not-so-n1086241. Accessed February 5, 2020. 10

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.

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.

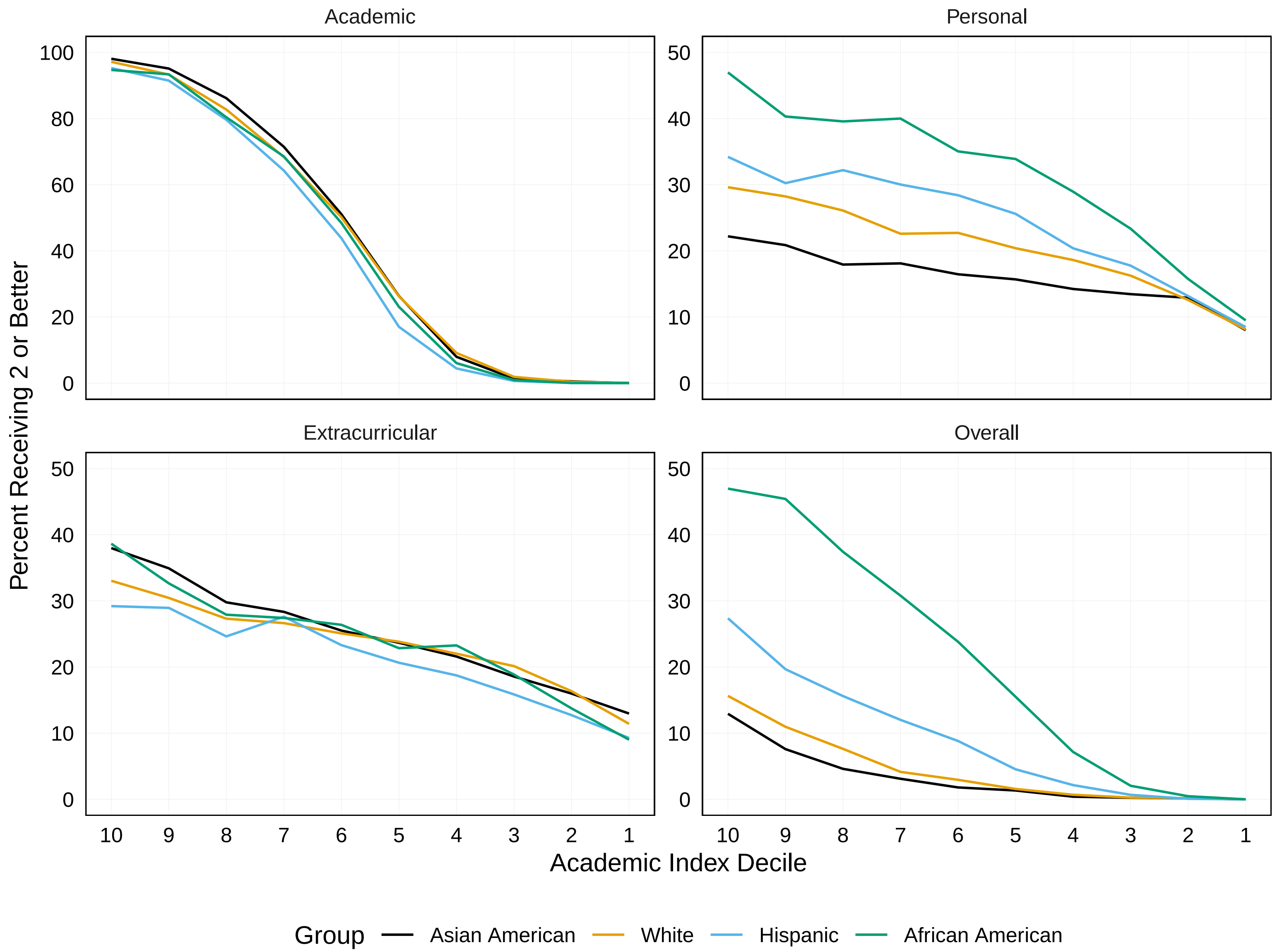
Table 3: Share of Applicants Receiving a 2 or Better on Application Ratings

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.

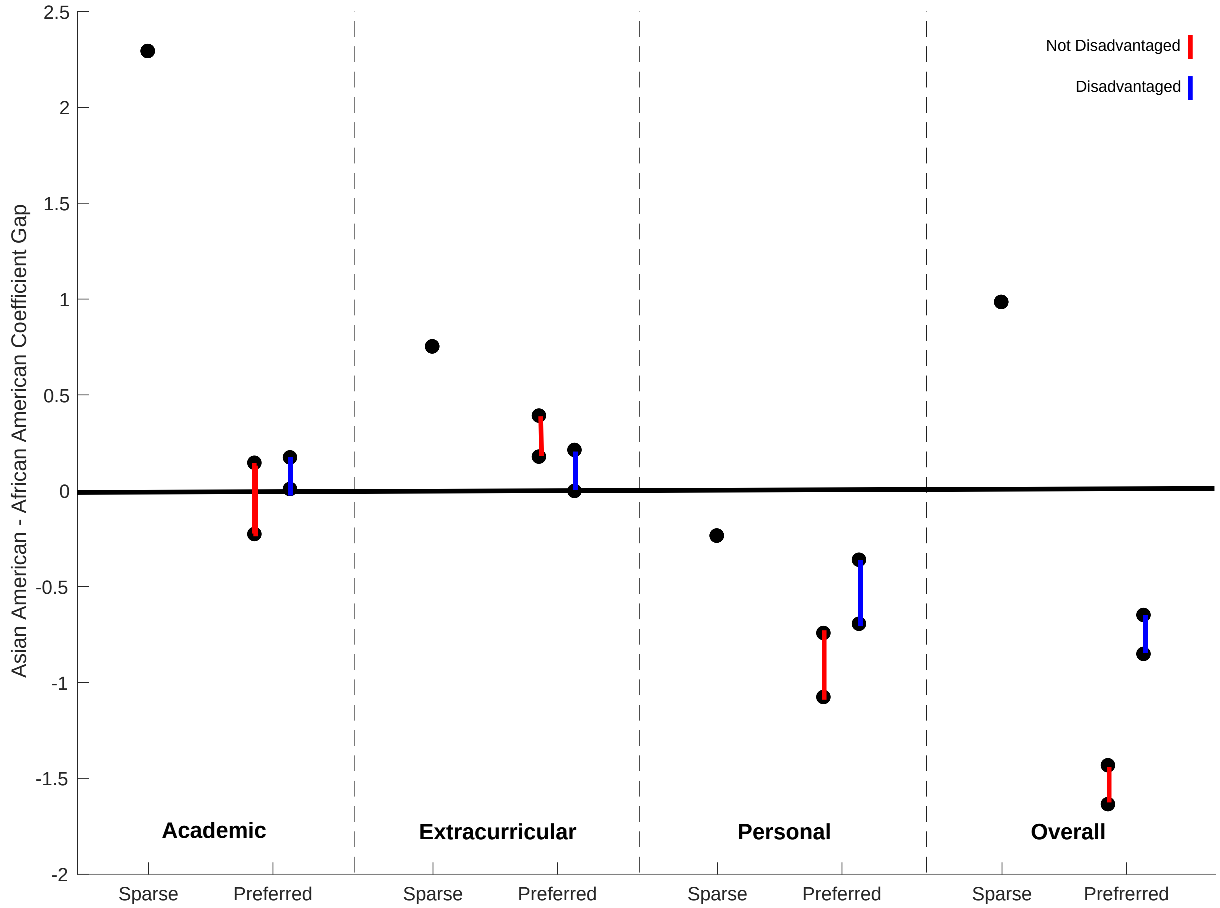
Table 4: Asian American Ratings and Admit Rate in Top Decile of Academic Index, Compared to Other Race Groups in Top Decile

		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.

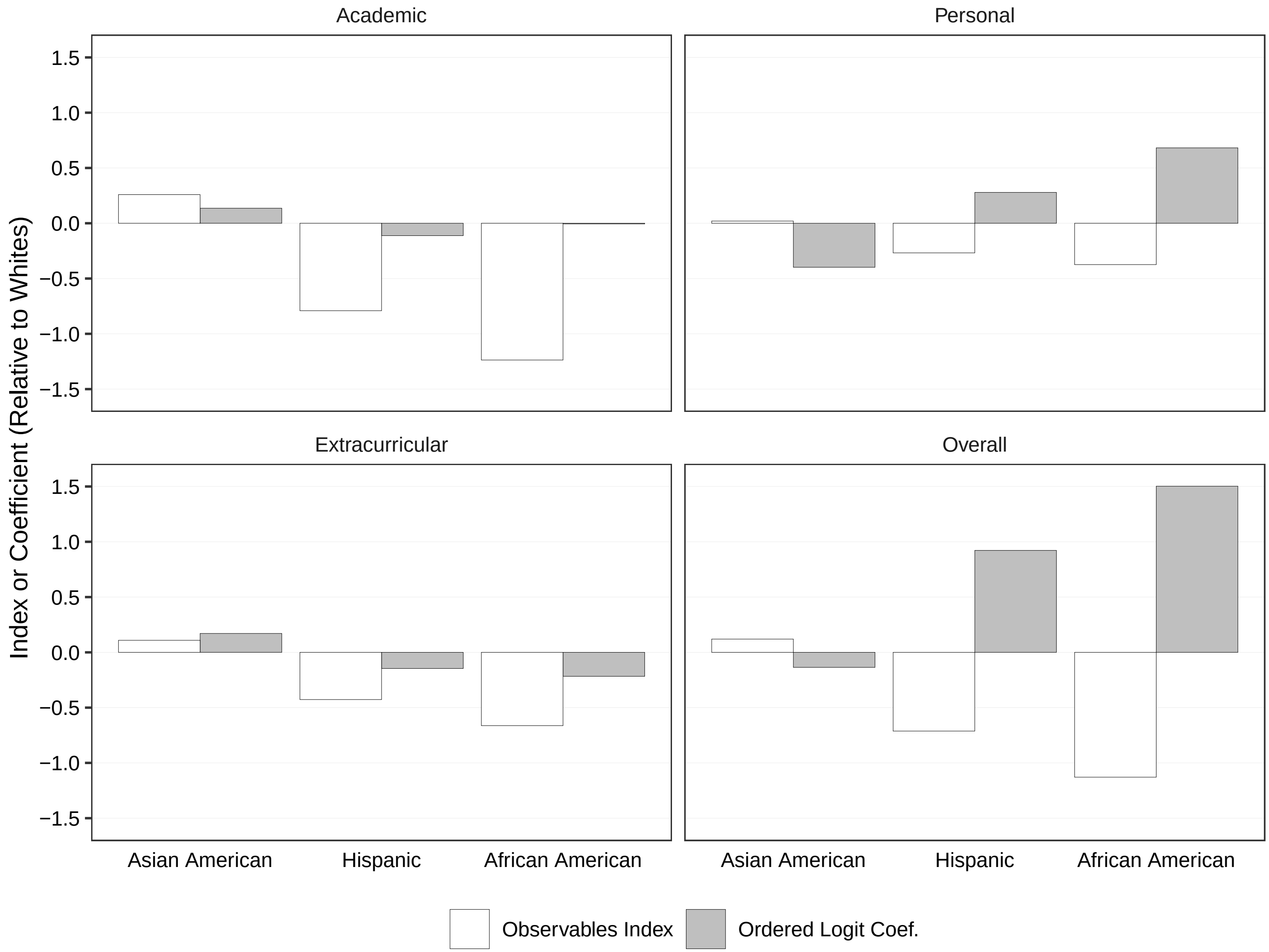
Figure 2: Estimated Ratings Gaps between Asian Americans and African Americans with Varying Number of Controls



Source: Authors' calculations from results reported in Appendix Tables D5 and D6.

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 3: 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 5: Selected Coefficients, Preferred Admissions Model

African American	3.772	Academic index	0.673	Disadvantaged × Year=2015	-0.041
	(0.105)		(0.189)		(0.157)
Hispanic	1.959	AI Sq. × (AI>0)	0.108	Disadvantaged × Year=2016	-0.263
	(0.085)		(0.106)		(0.165)
Asian American	-0.466	AI Sq. × (AI<0)	-1.236	Disadvantaged × Year=2017	-0.240
	(0.070)		(0.219)		(0.171)
Missing	-0.379	Female × Humanities	0.070	Disadvantaged × Year=2018	-0.496
	(0.122)		(0.110)		(0.175)
Year=2015	-0.875	Female × Biology	-0.132	Disadvantaged × Year=2019	-0.541
	(0.277)		(0.096)		(0.165)
Year=2016	-0.701	Female × Phys Sci	-0.101	Early × Year=2016	-0.304
	(0.287)		(0.136)		(0.128)
Year=2017	-0.528	Female × Engineering	0.062	Early × Year=2017	-0.171
	(0.290)		(0.111)		(0.126)
Year=2018	-1.102	Female × Math	0.028	Early × Year=2018	0.241
	(0.298)		(0.154)		(0.127)
Year=2019	-1.142	Female × Comp Sci	0.284	Academic Rating=4	-3.990
	(0.289)		(0.202)		(0.626)
Female	0.163	Female × Unspecified	0.215	Academic Rating=2	1.425
	(0.110)		(0.251)		(0.090)
Disadvantaged	1.660	Female × African Am.	-0.099	Academic Rating=1	4.094
	(0.138)		(0.114)		(0.156)
First generation	-0.014	Female × Hispanic	0.117	Extracurricular Rating=5	1.147
	(0.167)		(0.104)		(0.215)
Early Action	1.410	Female × Asian Am.	0.229	Extracurricular Rating=4	-1.301
	(0.104)		(0.082)		(0.393)
Waiver	0.697	Female × Missing	0.074	Extracurricular Rating=2	1.990
	(0.063)		(0.146)		(0.082)
Apply for Financial Aid	0.343	Disadv × African Am.	-1.577	Extracurricular Rating=1	4.232
	(0.101)		(0.143)		(0.169)
Humanities	0.206	Disadv × Hispanic	-0.582	Athletic Rating=5	0.761
	(0.137)		(0.133)		(0.087)
Biology	-0.031	Disadv × Asian Am.	0.144	Athletic Rating=4	-0.182
	(0.128)		(0.119)		(0.038)
Physical Sciences	0.027	Disadv × Missing	-0.035	Athletic Rating= 2	1.368
	(0.173)		(0.222)		(0.114)
Engineering	-0.082	Early × African Am.	0.039	>= 2 Acad. and Extra.	-0.143
	(0.149)		(0.152)		(0.088)
Mathematics	-0.276	Early × Hispanic	0.031	>= 2 Acad. and Athletic	-0.149
	(0.192)		(0.139)		(0.116)
Computer Science	0.191	Early × Asian Am.	-0.048	>= 2 Extra. and Athletic	-0.446
	(0.274)		(0.100)		(0.101)
Unspecified	-1.297	Early × Missing	0.192		
	(1.146)		(0.162)		

$N = 128,422$ & Pseudo $R^2 = 0.56$

Source: Model 5 in Table B.7.1R of Document 415-9.

Note: Standard errors in parentheses. The excluded categories are: white (race), 2014 (year), social sciences (major), and a rating of 3 for academic, extracurricular, and athletic ratings. A full list of controls is available in Appendix A.

Table 6: Inclusion of ALDC Applicants Distorts Effect of Other Admissions Criteria

	Typical	Typical & LDC	Typical & ALDC	% Increase with LDC	% Increase with ALDC
Academic Rating=4	-3.990	-2.426	-1.184	64.5%	237.1%
Academic Rating=2	1.425	1.206	1.209	18.2%	17.8%
Academic Rating=1	4.094	3.806	3.787	7.6%	8.1%
Extracurricular Rating=4	-1.301	-0.952	-0.171	36.7%	662.1%
Extracurricular Rating=2	1.990	1.689	1.646	17.8%	20.9%
Extracurricular Rating=1	4.232	3.795	3.726	11.5%	13.6%

Source: Data presented in Table 2 of Exhibit 287.

Notes: This table also appears in Arcidiacono, Kinsler, and Ransom (2019b) as Table 12.

Table 7: 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 (7).

Table 8: 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 applicants by their estimated admission index (i.e. the controls times their coefficients), absent race and admissions cycle.

Table 9: 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 applicants on the given dimension by their estimated admissions index, absent race. 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.

A 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

B The Robustness of the Asian American Penalty

The results from our preferred admissions model indicate large and statistically significant penalties against Asian American applicants. In this section, we investigate the robustness of this finding. More precisely, we show how the estimated Asian American penalty changes when we alter the set of control variables and move from a pooled to a yearly admissions model.

B.1 Additional Controls

In addition to excluding controls that are directly a function of race, we also exclude controls that are measured inconsistently or are already captured by other variables included the model. A prime example of this is parental occupation. In the database made available as part of the *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.[72] 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.[73] 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

[72] See Appendix Table D9 for evidence on the inconsistency of the parental occupation variables.

[73] 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."

being blue-collar.[74] For these reasons, 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.[75]

However, Document 419-141 advocates including both of these variables. For the occupation controls, Document 419-141 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.[76] When occupation and intended career are added to the preferred model, the marginal effect associated with Asian American is -0.75% and statistically significant at the 5% level.[77] While smaller than our preferred model estimates, it still indicates a large penalty for Asian American applicants relative to white applicants.[78]

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

[75] See Appendix Table D10 for information on how intended career varies by admissions cycle. Intended career is also never discussed in the reader guidelines for 2017 (Trial Exhibit DX 016), 2018 (Trial Exhibit P001), 2019 (Trial Exhibit P071) or 2023 (Trial Exhibit P633).

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

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

[78] Even if we include the personal rating, parental occupation, and intended career, the Asian American marginal effect is -0.34% and statistically significant at the 5% level (see Table 8.2N in Document 415-9). If Asian Americans were treated like similarly situated white applicants their admission probability would rise by 7%.

B.2 The Fragility of a No Discrimination Result

In Section B.1 we used our preferred specification as a baseline, and examined how changes from this baseline alter the estimated admissions penalty against Asian Americans. An alternative approach is to start from a baseline specification that makes a number of questionable modeling decisions to achieve 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, an alternative specification that finds no penalty is quite fragile.

B.2.1 Pooled Models

We begin by exploring the sensitivity of a pooled admissions model that is capable of generating a small and insignificant Asian American penalty. This is the pooled model advocated for in Document 419-141, and is quite different from our preferred specification. For the full details see Section 5 of Document 419-141. Here, we focus on versions of this model that exclude ALDC applicants, but it is important to note that Document 419-141 and Document 419-143 never estimate an admissions model excluding this special set. As we illustrated earlier in Section 3, it is inappropriate to include these applicants unless indicators for ALDC are interacted with all the other applicant attributes. A simpler approach is simply to exclude them.

The key differences between the pooled model in Document 419-141 and our preferred pooled specification are: (*i*) the exclusion of interactions between race and disadvantaged status; (*ii*) inclusion of the personal rating; and (*iii*) inclusion of parental occupation. While these three differences will be our primary focus, there are other relevant differences between the models. The pooled admissions model in Document 419-141 includes indicators for intended career and a full set of indicator variables for all profile ratings combinations and excludes interactions between race and gender, race and missing SAT II scores, and interactions between admissions cycle and disadvantaged, first-generation, early action, financial aid, permanent resident, intended major, flag for converted GPA=35, and missing SAT2 average. While we believe these additional differences are also potentially problematic, we do not study them in detail.

Appendix Table D11 shows how the estimated Asian American penalty is affected by the modeling choices associated with points (*i*), (*ii*), and (*iii*) above. 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. 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 the pooled model in Document 419-141 have a relatively minor impact.

While we have already discussed at length the concerns regarding the personal rating and parental occupation, it is worth discussing briefly why it is appropriate to interact race and disadvantaged status. An interaction between race and disadvantaged status makes sense when disadvantage has a different effect for different races. There is clear evidence that while Harvard gives African-American applicants a large preference, it does not give disadvantaged African-American students any preference for being disadvantaged (see Table 9 of Arcidiacono, Kinsler, and Ransom, 2019b). Thus, the effect of being disadvantaged is different across racial lines. So long as African Americans are used in the estimation of the model, the model requires these interaction terms. If the interaction terms are excluded, it significantly weakens the effect of disadvantage as an explanatory term. Since more Asian American applicants than white applicants are disadvantaged, the weaker effect of disadvantaged status in turn weakens the distinctions between white and Asian American applicants, thus tending to conceal the magnitude of discrimination against Asian American applicants.

B.2.2 Yearly Models

While Document 419-141 estimates a pooled admissions model, the preferred approach in Document 419-141 is one that estimates admissions preferences separately by year. 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 cycles.

One important difference between the pooled and yearly models in Document 419-141 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. 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). Document 419-141 defines a primary activity as an activity the applicant lists in the first or second activity field of her application. Additionally, the primary activities are collapsed into one of twelve groups in a somewhat arbitrary manner. More 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.[79] 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 in Document 419-141 at face value. Second, we define our own set of extracurricular controls. Rather than use the activity groupings proposed by Document 419-141, 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

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

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 the baseline yearly model in Document 419-141 is robust to: (*i*) the exclusion of interactions between race and disadvantaged status; (*ii*) inclusion of the personal rating; and (*iii*) inclusion of parental occupation. Appendix Table D12 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 in Document 419-141. 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 the yearly model from Document 419-141, but use the corrected extracurricular measures. In this case, the Asian American penalty is statistically significant when any of (*i*), (*ii*), or (*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 D12 mask important heterogeneity in the size and significance of the Asian American penalty across admissions cycles. In Appendix Table D13 we provide the year-by-year estimates of the Asian American penalty for the baseline specification in Document 419-141, 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 in Document 419-141. 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 D12).

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 one of these modeling choices does not alter the main finding.

C 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 in 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 C1, located at the end of this section, 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.[80] 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

[80] 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%.[81] 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%.[82]

[81] The accuracy rate for admits is calculated as the 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.

[82] 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 $94.55\% \times 94.55\% + 5.45\% \times 5.45\% = 90.05\%$.

Table C1: 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.

D Supporting Figures and Tables

Table D1: List of Legal Documents Used

Document	Description
Document 415-8	Plaintiff's expert witness opening report
Document 419-141	Defendant's expert witness opening report
Document 415-9	Plaintiff's expert witness rebuttal 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
Exhibit 287	Declaration of plaintiff's expert witness
Trial Exhibit DX 002	2012 Harvard admissions reader casebook
Trial Exhibit DX 016	Class of 2017 application reading procedures
Trial Exhibit DX 024	2012 Harvard admissions reader casebook discussion guide
Trial Exhibit DX 680	Primary extracurricular activity by race
Trial Exhibit P001	Class of 2018 application reading procedures
Trial Exhibit P071	Class of 2019 application reading procedures
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

Table D2: 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 (2019c) as Table 5.

Table D3: 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 (2019c) as Table 5.

Table D4: 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 D5: 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 D6: 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 D7: 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 D8: Selected Coefficients, Admissions Models

	(1)	(2)	(3)	(4)	(5)
African American	0.531	2.417	2.671	2.851	3.772
	(0.040)	(0.050)	(0.074)	(0.078)	(0.105)
Hispanic	0.425	1.273	1.286	1.339	1.959
	(0.039)	(0.044)	(0.063)	(0.067)	(0.085)
Asian American	0.057	-0.434	-0.565	-0.378	-0.466
	(0.032)	(0.035)	(0.052)	(0.055)	(0.070)
Missing	0.012	-0.283	-0.348	-0.330	-0.379
	(0.054)	(0.057)	(0.093)	(0.099)	(0.122)
Female	-0.044	0.254	0.228	0.271	0.163
	(0.025)	(0.027)	(0.064)	(0.088)	(0.110)
Disadvantaged	1.183	1.257	1.497	1.606	1.660
	(0.042)	(0.048)	(0.071)	(0.108)	(0.138)
First Generation	-0.004	0.174	0.161	-0.018	-0.014
	(0.052)	(0.059)	(0.059)	(0.127)	(0.167)
Early Action	1.616	1.456	1.371	1.348	1.410
	(0.032)	(0.035)	(0.055)	(0.084)	(0.104)
Waiver	-0.153	0.484	0.499	0.387	0.697
	(0.041)	(0.047)	(0.046)	(0.049)	(0.063)
Applied for Financial Aid	0.054	0.073	0.057	0.114	0.343
	(0.032)	(0.033)	(0.034)	(0.081)	(0.101)
Academic index		2.011	1.937	1.930	0.673
		(0.137)	(0.136)	(0.142)	(0.189)
AI Sq. $\times$ (AI>0)		0.268	0.339	0.390	0.108
		(0.078)	(0.079)	(0.082)	(0.106)
AI Sq. $\times$ (AI<0)		-0.950	-0.926	-0.933	-1.236
		(0.167)	(0.165)	(0.171)	(0.219)
N	142,728	142,700	142,700	136,061	128,422
Pseudo R Sq.	0.078	0.260	0.262	0.283	0.556
Demographics	Y	Y	Y	Y	Y
Academics	N	Y	Y	Y	Y
Race and Gender Interactions	N	N	Y	Y	Y
HS and NBHD Variables	N	N	N	Y	Y
Ratings (excluding Personal)	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.

Table D9: 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 D10: 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 D11: Is a Pooled Model that finds no Asian American Penalty Robust?

	Average Marginal Effect
(1) Baseline pooled model from Document 419-141	-0.22%
(2) Interact race and disadvantaged	-0.32%*
(3) Exclude personal rating	-0.65%*
(4) Exclude parental occupation	-0.37%*
(5) Combine (2), (3), & (4)	-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 D12: Is a Yearly Model that finds no Asian American Penalty Robust?

	Yearly with Document 419-141 extracurriculars	Yearly with corrected extracurriculars	Pooled
(1) Baseline model from Document 419-141	-0.18%	-0.24%	-0.22%
(2) Interact race and disadvantaged	-0.29%	-0.36%*	-0.32%*
(3) Exclude personal rating	-0.56%*	-0.62%*	-0.65%*
(4) Exclude parental occupation	-0.39%*	-0.47% *	-0.37%*
(5) Combine (2), (3), & (4)	-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 D13: Yearly estimates of the Asian American Penalty

	(1)	(2)	(3)	(4)	(5)
	Document 419-141 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.