NBER WORKING PAPER SERIES

LEGACY AND ATHLETE PREFERENCES AT HARVARD

Peter Arcidiacono
Josh Kinsler
Tyler Ransom

Working Paper 26316
http://www.nber.org/papers/w26316

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2019

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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Peter Arcidiacono, Josh Kinsler, and Tyler Ransom. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Legacy and Athlete Preferences at Harvard
Peter Arcidiacono, Josh Kinsler, and Tyler Ransom
NBER Working Paper No. 26316
September 2019
JEL No. I23,I24,J15

ABSTRACT

The lawsuit Students For Fair Admissions v. Harvard University provided an unprecedented look at how an elite school makes admissions decisions. Using publicly released reports, we examine the preferences Harvard gives for recruited athletes, legacies, those on the dean's interest list, and children of faculty and staff (ALDCs). Among white admits, over 43% are ALDC. Among admits who are African American, Asian American, and Hispanic, the share is less than 16% each. Our model of admissions shows that roughly three quarters of white ALDC admits would have been rejected if they had been treated as white non-ALDCs. Removing preferences for athletes and legacies would significantly alter the racial distribution of admitted students, with the share of white admits falling and all other groups rising or remaining unchanged.

Peter Arcidiacono
Department of Economics
201A Social Sciences Building
Duke University
Durham, NC 27708
and NBER
psarcidi@econ.duke.edu

Tyler Ransom
Department of Economics
University of Oklahoma
158 CCD1
308 Cate Center Drive
Norman, OK 73072
ransom@ou.edu

Josh Kinsler
University of Georgia
jkinsler@uga.edu

1 Introduction

The Students for Fair Admissions (SFFA) lawsuit against Harvard University provided unprecedented access to how Harvard makes admissions decisions and to the data underlying those decisions. While the focus of the lawsuit was on Asian American discrimination (relative to whites) and the size of racial preferences, the data provided in the lawsuit also revealed how preferences operate for other distinct applicant groups, including recruited athletes, legacies, those on the dean's interest list, and children of faculty and staff (ALDCs).[1] The aim of this paper is to provide a detailed analysis of Harvard ALDC applicants and their admissions outcomes relative to their non-ALDC peers.

Following on the heels of the recent college admissions scandal (Chappell and Kennedy, 2019), the treatment of ALDC students in college admissions is receiving renewed scrutiny. Wealth inequality in the US has been expanding for decades and college admissions preferences for applicant groups that may already be advantaged are generally condemned. Additionally, there are widespread concerns about the "fairness" of college admissions decisions. Applicants with greater academic preparation and accomplishments expect to be admitted at higher rates relative to less qualified applicants. Preferences for ALDC applicants have the potential to subvert this meritocratic ideal. The ongoing debate about the use of affirmative action in college admissions has also drawn attention to the impact that other admissions preferences have on the racial composition of college students. For example, it is widely accepted that legacy preferences will tend to benefit white applicants relative to other racial groups. Finally, underlying the rising concerns about privilege and fairness in college admissions is the growing competitiveness of the higher education market. Over the past twenty years, application levels have risen dramatically at elite colleges and universities

[1] The term ALDC was first used in the defendant's expert witness rebuttal report (Document 419-143, p. 30). As indicated in the Day 3 Trial Transcript, the dean's interest list contains a set of applicants that is of special importance to the dean of admissions. In particular, this list will include applicants whose parents have donated to Harvard, and applicants whose relatives have donated to Harvard. Trial Exhibit P104 and Trial Exhibit P106 illustrate how the dean's interest list operates. P104 is an email from the dean of the Kennedy School of Government at Harvard praising the admissions dean for admitting applicants whose families committed funding for buildings and fellowships prior to any admissions decision. P106 indicates a case where an applicant's score on the dean's interest list is below the maximum because the development office doesn't "see a significant opportunity for further major gifts."

in the US, with essentially no change in the number of seats available.[2]

Harvard epitomizes the competitive nature of elite college admissions—the admit rate for the Class of 2023 was 4.5%—and is thus an ideal setting to explore the attributes of ALDC applicants and their admissions outcomes relative to their non-ALDC peers.[3] Harvard is the oldest university in the US and is considered to be one of the most, if not the most, preeminent educational institution for societal leaders.[4] While of interest in its own right, Harvard is also similar to other Ivy League and highly selective institutions in terms of academics and the socioeconomic background of its students.[5]

While detailed data on admissions are typically tightly guarded by universities, rich data on Harvard's decisions were made available as a result of the court case. The Harvard admissions data contain a plethora of applicant characteristics, including detailed information on demographics, academics, and extracurricular activities. Of particular importance, the data contain information on Harvard's internal ratings of its applicants on a host of dimensions, including academic, extracurricular, athletic, and personal ratings. The availability of Harvard's internal ratings allows us to better describe the differences between ALDC and non-ALDC applicants, as well as to account for factors that would typically be unavailable when estimating admissions models.

We use the detailed Harvard admissions data to examine three questions:

1. How do the characteristics of recruited athletes, legacies, and other special groups of applicants and admits compare to the corresponding characteristics of the broader applicant and admit pools?

2. How large of an admissions advantage is given to those in special recruiting categories?

3. How would the removal of these admissions preferences impact legacy and athlete

[2]See Bound, Hershbein, and Long (2009) and DeSilver (2019).

[3]Throughout, class refers to the year Harvard would expect the applicant to graduate college, i.e. four years after high school graduation.

[4]Five US presidents and ten Supreme Court justices attended Harvard as undergrads, the most of any institution in the US.

[5]All eight Ivy League universities are ranked in the top twenty national universities in the 2019 US News and World Report (`https://www.usnews.com/best-colleges/rankings/national-universities`). Chetty et al. (2017b) show that Harvard has a similar share of students from the top 1% and top 20% of the US income distribution as compared to the other seven Ivy League institutions plus MIT, Stanford, Chicago, and Duke.

admissions and the racial composition of admitted students?

These questions are motivated by the ongoing public debate related to privilege, fairness, and racial equity in higher education.

In response to the first question, we find that for each special applicant group under the ALDC umbrella, applicants and admits are disproportionately white and come from higher income households. As an example, 40% of non-ALDC applicants are white, while nearly 70% of legacy applicants are white. When we explore other characteristics, such as academic preparation, extracurricular strength, and personal qualities, the results are more nuanced. On average, LDC applicants (that is, excluding athletes) are stronger than non-ALDC applicants. However, the average LDC admit is weaker than the average non-ALDC admit, suggesting an admissions advantage for LDC applicants. The admissions advantage for recruited athletes appears to be even stronger. Admitted athletes have significantly worse credentials than non-ALDC admits, and in some cases, non-ALDC *applicants*.[6]

We show that LDC applicants are especially highly rated on Harvard's non-academic ratings. Ratings related to extracurricular activities, athletics, and personal qualities are included in the admissions process to allow for a more comprehensive view of an applicant. It is widely believed that this "holistic" approach—pursued by many competitive universities—opens doors for less advantaged applicants. Harvard's non-academic ratings appear to achieve that goal only if race is the central characteristic defining disadvantage: the gap between white and under-represented minority applicants is much smaller for non-academic ratings than for the academic rating. However, *within* racial groups, these holistic admissions criteria favor advantaged applicants. As a result, LDC applicants within racial groups are doubly advantaged in the admissions process. They benefit from the particular criteria Harvard uses to evaluate applicants, and they receive an additional tip in admissions decisions conditional on those criteria. These findings are consistent with Lee and Ries (2019), who find that use of non-academic admissions criteria favors white and wealthy applicants at the University of British Columbia.

[6]One might suppose that recruited athletes are a small share of admitted students to Harvard. They are not, representing 10% of admits. In fact, Harvard offers 42 Division I intercollegiate sports teams—the most in the nation. See https://college.harvard.edu/student-residential-life/athletics. For more information about Harvard athletics, see Appendix E. In contrast, Ohio State University, one of the largest public universities in the country, fields only 37 Division I teams. See ohiostatebuckeyes.com.

To address our second question, we estimate a model of Harvard's admissions that accounts for hundreds of applicant characteristics, including Harvard's academic, extracurricular, and athletic ratings, among others. Admissions preferences for ALDC applicants are substantial. We find that a white non-ALDC applicant with a 10% chance of admission would see a five-fold increase in admissions likelihood if they were a legacy; more than a seven-fold increase if they were on the dean's interest list; and that they would be admitted with near certainty if they were a recruited athlete.[7]

Finally, we explore how the admitted class at Harvard would change if ALDC preferences were eliminated. First, we estimate that only one quarter of white ALDC admits would have been admitted had they been treated as white non-ALDC applicants. Given the highly advantaged status of this group, eliminating ALDC preferences would tend to reduce the household income level among Harvard admits. Second, we explore how the number of admits in each racial/ethnic group would change if legacy and athlete preferences were removed, holding fixed the total number of admits. We find that removing either of these preferences would result in significantly fewer white admits with increases or no change in the number of African American, Hispanic, and Asian American admits.[8]

A natural question is why Harvard would give such large preferences to these groups. In a recent report issued by Harvard on race-neutral admissions (Trial Exhibit P316), a committee of Harvard deans provides various rationales for promoting ALDC preferences. While we forego the details of these arguments, the main takeaway is that these preferences are thought to build a sense of community. Harvard's argument is that students who are currently on campus benefit from being exposed to recruited athlete and legacy students, while the university benefits from increased engagement and donations from alumni.[9] There are other incentives for Harvard to employ these preferences as well. For example, donations

[7]We focus on whites because they make up the vast majority of ALDC applicants and admits.

[8]Removing preferences for recruited athletes leaves the number of African Americans essentially unchanged, with increases for Hispanic and Asian American admits. Removing legacy preferences increases the number of admits for each of the non-white groups.

[9]A major component of alumni engagement is Harvard's 10,000+ alumni that volunteer to interview and recruit applicants (Trial Exhibit P316, pp. 9, 16–17). A higher admissions probability for the children of these alumni implicitly compensates them for their contribution to the admissions process. When considering ramifications of removing legacy preferences, Harvard said it is "concerned that [doing so] would diminish this vital sense of engagement and support" (ibid., p. 17). These arguments are also mentioned in Document 421-9, pp. 191–192.

by alumni—as opposed to parents of current students—are part of the formula used in *U.S. News & World Report* Best Colleges rankings,[10] providing an additional incentive to admit legacies whose parents will be more likely to donate if their child is admitted.

While there is little research on the peer effects of athletes and legacies, researchers have analyzed the impact that athletics and legacy admissions have on giving. Meer and Rosen (2009b) and Anderson (2017) show that a university's athletic success can lead to increased donations and prestige. Meer and Rosen (2009a), Meer and Rosen (2010), and Clotfelter (2003) find an increased propensity to donate to a university among alumni parents and students who graduate as legacies. So while the general public may find ALDC preferences unseemly, institutions see them as a way of maintaining and enhancing their prestige (Jacob, McCall, and Stange, 2018).[11]

Due to the paucity of admissions data, the number of papers analyzing legacy and athlete admission preferences is limited. Espenshade, Chung, and Walling (2004) use admissions data from three elite research universities to estimate the admissions tip that legacies and athletes receive conditional on SAT scores, race, and gender. They find that legacy and athlete status increases the odds of admission by three and four, respectively. In our preferred model, the similar increases in odds are substantially higher at over eight (legacies) and *five thousand* (athletes) times. This reflects at least two factors: our model provides substantially more explanatory power than their models due to the wealth of data provided;[12] and there is evidence that legacy and athlete preferences have been increasing over time.[13] Hurwitz

[10]The rankings methodology of *U.S. News & World Report* is described at `https://www.usnews.com/education/best-colleges/articles/how-us-news-calculated-the-rankings`.

[11]Preferences for ALDC applicants can also be rationalized if the economic returns to attending an elite institution are smaller for students from less advantaged backgrounds. There is conflicting evidence regarding this question. Zimmerman (2019) and Riehl (2018) find that the return to attending a high quality college is larger for students from wealthy backgrounds. On the other hand, Dale and Krueger (2014) and Zimmerman (2014) find that the returns to college quality are large for low-income and historically disadvantaged groups.

[12]The Pseudo R^2 of their models are around 0.2 while our preferred model has a Pseudo R^2 of 0.56. The more explanatory power of the model, the higher the log-odds all else equal, due to the coefficient estimates of logit models being estimated relative to the variance of the unobservables. The more explanatory power of the model, the lower the variance of the unobservables. See Norton and Dowd (2018) for a discussion of this issue. While the odds ratio of 5,000 may seem ridiculous, this is simply the exponentiated logit coefficient, where $\exp(8.532) = 5,075$, illustrating the massive advantages recruited athletes have in the admissions process.

[13]Arcidiacono, Kinsler, and Ransom (2019) demonstrate that the admissions advantages athlete and legacy applicants receive at Harvard have increased substantially over the past twenty years. In particular, the admit rate for athlete and legacy domestic applicants relative to the admit rate for non-athlete and non-legacy

(2011) uses data from thirty private colleges and universities and finds that legacy applicants are again three times more likely to be admitted. To help account for unobserved differences between legacy and non-legacy applicants, he exploits multiple applications per applicant in a fixed-effects-type model. An assumption of this model is that different schools value applicant attributes equivalently. Our approach instead uses more detailed applicant data and a single school's own set of internal ratings of applicants to help control for differences between ALDC and non-ALDC applicants. Despite the differences in approaches, our results are similar to his findings for legacies at the most selective schools in his data.

The rest of the paper proceeds as follows. In section 2, we provide background on Harvard's admissions process and the data we use. Section 3 compares the characteristics of ALDC applicants to their non-ALDC counterparts. Section 4 describes the admissions model and shows what the estimated preferences imply about how these advantages affect the admissions probabilities for particular applicants. Section 5 shows how removing legacy and athlete preferences would affect the racial distribution of the class, taking into account that removing these preferences makes it slightly easier for other applicants to be admitted. Section 6 concludes.

2 Harvard Admissions and Data

In this section, we describe the admissions process at Harvard and the key documents and methodology used to study ALDC preferences. All documents are publicly available either at the URL in the bibliography, or on the Public Access to Court Electronic Records (PACER) website at https://www.pacer.gov/.

2.1 Harvard Admissions Process

To provide a framework for understanding the role of ALDC preferences in Harvard admissions, we first describe the Harvard admission process.[14] As sources, we use the unsealed (but heavily redacted) depositions of Admissions Dean William R. Fitzsimmons (Document

domestic applicants has increased from a ratio of 4:1 to 9:1 between the Classes of 2000 and 2017.

[14]Harvard's process is similar to that employed by other Ivy League schools (Jackson, 2017; Sabky, 2017), and likely other elite institutions that employ holistic admissions criteria.

421-9) and Admissions Director Marlyn McGrath (Document 419-1), as well as two versions of the admissions office's reading procedures that were admitted to the court as Trial Exhibits P001 and P723. We also use trial testimony of Harvard admissions office personnel and exhibits filed as part of various legal motions.

2.1.1 Overview

Domestic applicants are first divided into "dockets" which correspond to geographical areas of the United States.[15] Each application is read by at least one admissions officer who assigns scores on a series of internal Harvard ratings, which we detail below.[16] In addition to the internal ratings produced by the admissions officer, most applicants are also interviewed and rated by a Harvard alum, and a very small number (less than 3%) may be interviewed by an admissions staff member.[17] The alumni and staff evaluate applicants on a subset of the application ratings, using the same scale that admissions officers employ.[18] Competitive applications may also be read by an additional admissions office member (typically the head of the docket sub-committee).[19] The initial reader of the application is known as the "first reader," whereas the docket head is known as the "third reader."[20]

Application scores, supporting material, and other characteristics of the applicant (such as demographics) are summarized in a single document known as the "summary sheet," which is used to convey application information to admissions officers.[21] Throughout the application cycle, admissions officers assigned within dockets meet together to discuss candidates and make tentative admissions decisions.[22] Towards the end of the cycle, the full committee (i.e. all admissions officers) meets to consider docket sub-committee recommen-

[15] See pp. 176–177 of Document 419-1

[16] See Trial Exhibit P001, p. 5. Those who wish to know more about the reading process can find full details in the casebook (Trial Exhibit DX 002) and its discussion guide (Trial Exhibit DX 024).

[17] See Trial Exhibit P001, p. 12; Document 419-1, p. 55; and Document 421-9, p. 191. Staff interviews are scheduled on a first-come, first-serve basis (Document 419-1, p. 96). Trial Exhibit P619 describes the frequency of staff interviews.

[18] See Document 421-9, p. 239.

[19] See Document 419-1, p. 173.

[20] ibid., pp. 172–178. A "second reader" generally applies to situations where the first reader is in training.

[21] See Document 419-1, pp. 96–97. Document 421-9 mentions that applicant demographics (such as race) are discussed in committee meetings (see p. 288), as well as disadvantaged and legacy status (see pp. 201, 259).

[22] See Document 419-1, pp. 184–185

dations and make final decisions.[23]

2.1.2 Scoring of Applications

A key advantage of using Harvard's data to study ALDC preferences is the availability of internal ratings for each applicant. In this section, we describe precisely how applications are scored.

Readers assign each application a set of numerical codes indicating its strength.[24] Readers give an overall rating as well as a rating profile, which is composed of ratings in the following areas: academic, extracurricular, athletic, and personal.[25] Additionally, competitive applications are rated on the "full profile," which includes the strength of support from the applicant's teachers and school counselor (school support ratings) and ratings assigned by alumni or staff interviewers.[26] Each of the ratings is on a five-point scale, with lower numbers indicating better ratings, and "+" and "−" suffixes respectively indicating better and worse ratings within a number.[27] Each of the component ratings, as well as the overall rating, are included on the summary sheet.[28] The overall ratings assigned by the first and third readers both appear on the summary sheet, but the other component rating scores are overwritten in the event of an additional read.[29]

We now highlight some key aspects of Harvard's reading procedures, which explain how to code numerical values for each of the ratings.[30] Our highlights below focus on the procedures used for the Class of 2018, because they correspond to the time period of our data. The reading procedures were recently updated for the Class of 2023. Harvard made a number of major changes, foremost being changes in the guidelines of how to score the personal rating.

[23] ibid., pp. 186–187; Document 421-9, pp. 287–290.
[24] See Trial Exhibit P001, p. 5.
[25] ibid.
[26] ibid.
[27] See Document 419-1, pp. 158–161, and Trial Exhibit P001, p. 5. For example, a rating of 1− is better than a 2+.
[28] See Trial Exhibit P001, p. 5.
[29] ibid.
[30] In addition to these reading procedures, Harvard provides new admissions officers and new alumni interviewers specific training on how to convert an application's data into a numerical score (Day 6 Trial Transcript, pp. 151–154). The backbone of this training is the case book (Trial Exhibit DX 002) and its discussion guide (Trial Exhibit DX 024).

Overall rating A 1 in this rating category is reserved for "exceptional" applications with "90+% [chance of] admission." On the other hand, a 4 is "neutral," while a 5 is "negative" and characterizes those candidates whose "credentials are generally below those of other candidates" (Trial Exhibit P001, p. 5). The overall rating is based on each of the component ratings (Document 419-1, p. 159) as well as any other factors that admissions officers deem valuable to Harvard, such as race or legacy or disadvantaged status.[31]

Academic rating The academic rating ranges from 1s who have "summa potential ... near-perfect [test] scores and grades" to 5s who have "marginal potential ... modest grades and 500 scores (25 and below ACT)" (ibid., pp. 5–6).

Extracurricular rating This rating also includes "community employment [and] family commitments." A 1 indicates "possible national-level achievement or professional experience ... truly unusual achievement." A 4 indicates "little or no participation" (ibid. p. 6).

Athletic rating A score of 1 on the athletic rating indicates a recruited athlete.[32] A 2 indicates "strong secondary school contribution in one or more areas; possible leadership role(s)." A 4 indicates "little or no interest" (Trial Exhibit P001, p.6) but is "not necessarily negative" (Document 419-1, p. 159).

Personal rating The personal rating criteria, compared to the other ratings, are surprisingly vague. The criteria are as follows: "(1) Outstanding; (2) Very strong; (3) Generally positive; (4) Bland or somewhat negative or immature; (5) Questionable personal qualities; (6) Worrisome personal qualities" (Trial Exhibit P001, pp. 6–7).[33]

School support rating Readers examine the letters of recommendation provided by the applicant's high school and code them on a range from "'The best ever,' 'one of the best in x years,' truly over the top" (1) to "Negative or worrisome report" (5).

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

[32] See Trial Exhibit P001, p. 6, and Document 419-1, pp. 160, 163.

[33] These criteria were substantially updated in Trial Exhibit P723. It has half of a page of contextual description leading into the description of the scoring criteria, and the criteria for scores of 1 or 2 each contain five or six sentences.

2.1.3 The Treatment of ALDCs in the Admissions Process

The testimony and documents made available at trial also provide an in-depth look at how ALDC applications are handled by the admissions staff. While ALDC applicants go through the full committee process with all other applicants, we find that ALDC applicants are given special treatment in the reading process. ALDC applications are closely monitored throughout the cycle by the admissions dean, admissions director, athletic coaches, and others.[34] Often, ALDC applications are annotated in specific ways to aid in the monitoring process.[35] As a group, ALDCs are also about 20 times more likely to interview with a member of the admissions office.[36]

The specialized treatment ALDC applicants receive likely contributes to their admissions advantages. Much of Harvard's specialized treatment of these applicants stems from its desire to find "distinguishing excellences," or individuals who are exceptional in at least one of a variety of dimensions (Trial Exhibit DX 005, pp. 9–10). Admissions "tips" are given to those applicants who have distinguishing excellence, and listed among these are athletic ability and legacy status (Trial Exhibit DX 005, pp. 9–11).[37] The rest of this section illustrates how each of the ALDC categories intersects with the admissions process.

Athletes While Harvard does not offer athletic scholarships,[38] each of its 42 sports teams has a liaison that moderates contact throughout the admissions process between the admissions office and that team's coach.[39] Additionally, the admissions office sets aside a certain number of staff interview slots for only athletes, which can occur outside of the time frame for which staff interviews are made available to the general public.[40]

[34] See below for specific references.

[35] See below for specific instances.

[36] See Trial Exhibit P619 for a summary of staff interview rates by ALDC status.

[37] Other specific examples of distinguishing excellence include the following: intellectual ability, personal qualities, capacity for leadership, creative ability, and socioeconomic background (including race/ethnicity). See pp. 9–11 of Trial Exhibit DX 005 for further details.

[38] Document 419-1, p. 160

[39] See Document 419-1, p. 41. See also footnote 89 of Document 419-143, which relates an email correspondence between women's hockey coach Katey Stone and the admissions office in which Stone advocates for her recruits. Stone admits that "their testing may not be that of others, yet what they will bring to the Harvard classroom, athletic area and community is immeasurable."

[40] See Day 5 Trial Transcript, p. 184.

Legacies Legacies are coded in the "Lineage" field of the application, meaning that legacies are strictly children of alumni, and not other relatives (Trial Exhibit P001, p. 3).[41] The reading procedures instruct that these files "should be read by [Admissions Dean Fitzsimmons] following the normal reading process if the decision might require special handling or if another reading might be helpful" (Trial Exhibit P001, p. 3).

Dean's Interest List There is no mention of the Admissions Dean's or Admissions Director's interest lists in the reading procedures. However, Document 421-9 provides a number of details about the handling of such applications. First, members of this list receive an additional rating which is separate from the profile ratings (ibid., pp. 269–273). This additional rating is tied to the applicant's (or the family of the applicant's) donation history and future donation prospects (ibid., p. 278).[42] When subcommittees discuss applicants on this list, the admissions dean may pre-emptively join the meeting to discuss members of this list, or may have individually discussed the applicant with the subcommittee chair beforehand (ibid., p. 279). A similar process applies for the Admissions Director's interest list (ibid., p. 278). Finally, being a member of the Dean's or Director's interest lists is a field included in the individual-level applicant data that were produced by Harvard for the SFFA case.

Children of Faculty or Staff The reading procedures instruct readers that these applications "should be sent to [Admissions Dean Fitzsimmons] after the normal reading process has been completed."[43]

2.2 Admissions Data

Our analysis of ALDC applicants and admissions is based upon anonymized data on domestic applicants produced by Harvard. The applicants come from the set of students who would be

[41] Klein and Keto (2015) show that 16.8% of the Class of 2019 are legacies by this strict definition, but that nearly 28% of the class is a legacy by a looser definition (i.e. parent, sibling, grandparent, aunt/uncle, or other relative graduated from Harvard). While these other definitions do not appear to be formally tracked by the admissions office, they may be correlated with other ALDC categories.

[42] See also Trial Exhibit P106 which alludes to this rating.

[43] Harvard labels some faculty as "staff." Only Faculty of Arts and Sciences are called "faculty," whereas faculty of other colleges are called "staff," in addition to non-faculty employees (Trial Exhibit P001, p. 3). This distinction motivates our decision to put faculty and staff children in the same category.

graduating from Harvard in 2014–2019. This range includes students who would typically have been applying to Harvard in the fall of 2009–2014 and graduating from high school in the spring of 2010–2015. In all, the sample consists of 166,727 domestic, non-transfer, complete applications.[44]

For each applicant, the data contain detailed demographic information, academic performance in a variety of categories, scores on each internal rating, and final admission decisions. Critically, the data also include indicator variables for each of the LDC categories. Recruited athletes are identified as those applicants who receive a 1 on the athletic rating.

It is important to point out that we no longer have access to Harvard's individual-level applicant data. As a result, the findings presented in the current paper are based solely on information in the publicly released versions of the expert witness reports or information publicly released in other documents. A full list of the documents we rely on is presented in Appendix A. Fortunately, the publicly available documents provide enough detailed information for us to infer the characteristics of ALDC applicants relative to their non-ALDC peers, and the preferences afforded to ALDC applicants in the admissions process. Appendix B provides both a general and detailed account of how each number we present is generated based on information in the public record.

3 Characteristics of ALDC Applicants and Admits

We begin by describing differences in admit rates and racial composition across ALDC status. We follow this with separate analyses of the ratings and attributes of LDC applicants and athlete applicants, since we are able to provide more information in greater detail for LDC applicants.[45]

[44] For further details about the data and sample selection, see sections 2.2 and 2.3 of Document 415-8 and section 3 of Document 419-141.

[45] Our analysis of athletes is limited because it is more difficult to isolate the attributes of recruited athletes in publicly available documents. Further details are provided in Appendix B.2.

3.1 ALDC admit rates

ALDC applicants are admitted at substantially higher rates than non-ALDC applicants. Table 1 illustrates this phenomenon.[46] It shows domestic admit rates for each of the ALDC groups across the six admissions cycles spanning 2014–2019. With an admit rate of 86%, recruited athletes are over 14 times as likely to be admitted as those that are not recruited athletes. This is by far the largest admit rate. Although recruited athletes are less than 1% of the applicant pool, they make up over 10% of the admitted class.[47]

The admit rate for legacy applicants over this period was 33.6%, which is 5.7 times higher than the admit rate for non-legacy applicants. Legacies are the largest of the ALDCs, both in terms of number of applicants as well as number of admits, and make up 14% of the admitted class. Interestingly, this legacy admit rate is similar to the legacy admit rate when the Office for Civil Rights (OCR) was investigating Harvard, despite Harvard receiving many more applications in recent years. In the last two admissions cycles that OCR analyzed (1991 and 1992), the admit rates for legacies ranged from 33.4% to 35.2%.[48]

Children of faculty and staff and those on the dean's interest list also have high admit rates at 46.7% and 42.7%, respectively, making up 1.3% and 9.5% of admitted students. Note that the total share of admits that belong to one of the ALDC groups is 29%, which is the less than the sum of the shares listed here.[49] The difference arises because of double counting. For example, some recruited athletes may also be legacies.

To get an initial sense of the admissions advantage these groups receive, Table 2 reports admit rates by academic rating for three sets of applicants: non-ALDC, LDC, and recruited

[46]The sample used to create Table 1 includes applicants that were subsequently dropped in the main analysis sample of Document 415-9 as the result of missing ratings, SAT scores, or academic index. Thus, the numbers in Table 1 will not necessarily match the numbers in other documents that use the trimmed sample. Note that the total number of applicants in Table 1 is 166,727, which matches the number of applicants from Table A.5 in Document 415-8 after the first three sample cuts. As a result, we know that the total number of admits in this sample is 11,132, which we can derive using Tables A.5R and B.3.2R in Document 415-9.

[47]The 10% number is calculated by multiplying the admit rate times the number of applicants and dividing by 11,132, the total number of admits. Similar calculations are used to generate admit shares for legacies, dean's list, and children of faculty/staff.

[48]See Trial Exhibit P555. OCR investigated whether Harvard discriminated against Asian Americans in its undergraduate admissions process, concluding that Harvard did not discriminate against Asian American applicants.

[49]The 29% figure is based on information from Table B.3.2 in Document 415-8 and Tables B.3.1R and B.3.2R in Document 415-9. This figure is also directly reported in Trial Exhibit DX 706.

athletes. The first column shows the results for non-ALDCs. Consistent with Harvard valuing academics, higher academic ratings are associated with higher admission rates. All non-ALDC applicants who received a 5 were rejected and those who received a 4 were almost certain rejects, with an admit rate of 0.02%. The bulk of non-ALDC Harvard applicants receive a 3 or a 2 on the academic rating and see average admit rates of 2.4% and 10%, respectively. Receiving a 1 is rare (less than 0.5% of applicants), but is rewarded with an admit rate of over 66%.

At ratings of 4 or better, LDC applicants have substantially higher admit rates than their non-ALDC counterparts (no one outside of athletes is admitted with an academic rating of 5). While a 4 virtually guarantees rejection for non-ALDC applicants, LDC applicants are admitted at a rate of 3.5%, which is a full percentage point higher than the admit rate for non-ALDC applicants who received a 3 (one rating better). LDC applicants who receive a 3 have admit rates of 18%, which is 7.5 times the admit rate of non-ALDC applicants who receive a 3 and 1.8 times higher than the admit rate for non-ALDC applicants who received a 2.

While the differences in admit rates between non-ALDC and LDC applicants conditional on the academic rating are striking, they pale in comparison to the differences between recruited athletes and everyone else. An academic rating of 5 guarantees rejection for all other applicants, but half of the recruited athletes with the worst academic rating are admitted. For all other academic ratings, the admit rate for recruited athletes was 79% or higher. Recruited athletes who receive a 4 on the academic rating had admit rates that were higher than non-ALDC applicants who received a 1.

3.2 The racial composition of ALDCs

The overall admit rates, as well as the admit rates conditional on the academic rating, suggest the possibility of large preferences for ALDC applicants. Panel A of Table 3 shows large admissions advantages for ALDC applicants within each racial group. However, this does not imply that all racial groups benefit evenly from ALDC preferences, since the share of each racial group that is ALDC varies considerably.

Panels B and C of Table 3 show the racial distribution of applicants and admits for non-

ALDCs as well as for each of the ALDC categories. The share of applicants and admits who are white in each of the ALDC categories are both substantially higher than the corresponding shares of applicants and admits who are white and not ALDC. For example, recruited athletes, legacies, and dean's interest list applicants are all over 68% white, yet the share of non-ALDC applicants who are white is less than 41%. All other racial groups see higher representation among non-ALDC applicants and admits than in any of the corresponding ALDC applicant and admit categories.[50]

The final panel of Table 3 documents the share of admits who are non-ALDC and the share of admits in each of the ALDC categories by race. For non-whites, the share of ALDC admits are less than 16% for each racial/ethnic group.[51] The corresponding share for whites is much higher at over 43%. Indeed, the share of white admits who are recruited athletes *alone* is higher than the share of ALDC admits *as a whole* for any of the other racial/ethnic groups.

3.3 Comparing LDC and non-ALDC Applicants and Admits

The previous section shows that ALDC applicants are admitted at much higher rates and are substantially more likely to be white than non-ALDC applicants. In this section, we examine further how LDC applicants differ from non-ALDC applicants. We focus on three types of measures: background characteristics, Harvard internal ratings, and academic preparation. Since we are unable to fully replicate this analysis for recruited athletes, we return to recruited athlete attributes in the next subsection.

3.3.1 Demographics

Table 4 displays the background characteristics of LDC and non-ALDC applicants by race and admission status. Panel A lists these statistics for all applicants and Panel B for all admits. The results in Panel A indicate that, among applicants of each race, LDCs are significantly less likely to be coded as disadvantaged or to receive financial aid. They are

[50] There is one exception: Asian Americans represent a slightly greater share of children of faculty and staff admits than non-ALDC admits. However, this category is much smaller than the other ALDC components.

[51] Because of overlap in ALDC group membership, we compute this as 100 minus the non-ALDC share.

also much more likely to have their application reviewed by a second admissions officer.[52] The increased propensity to obtain a second review is consistent with Harvard's reader guidelines and handling of special applications as detailed in Section 2.1.

Panel B of Table 4 shows that admitted LDCs also come from more privileged backgrounds than non-ALDC admits. We investigate the relative advantage of legacy admits further by using publicly available survey data on Harvard enrollees. *The Harvard Crimson*, the nation's oldest continuously published daily college newspaper, surveys each incoming class about their family background. These surveys confirm the advantaged status of legacies. For the Class of 2019, 40.7% of legacy respondents have parents who earn more than $500K, which is the Top 1% of US income.[53] The corresponding share for all respondents, including legacies, is only 15.4%.[54]

Finally, comparing across Panels A and B also highlights that disadvantaged status affects admissions probabilities differently for non-ALDC and LDC applicants. Among non-ALDCs, disadvantaged applicants tend to be over-represented among admits relative to all applicants. The reverse is true for LDCs, suggesting that applicant characteristics are valued differently according to LDC status.

3.3.2 Harvard Ratings

In addition to differences in basic applicant characteristics, we are also able to explore whether Harvard rates LDC applicants differently than non-ALDC applicants. We measure applicant strength as the rate of receiving a score of 2 or better on Harvard's internal

[52] Recall from Section 2.1 that Harvard refers to an additional reader as the "third" reader.

[53] See Klein and Keto (2015). The share of legacies whose parents earn above $500K is not reported directly. However, the survey reports the probability of being a legacy, the distribution of family income for all respondents, and legacy by family income. We use these numbers to calculate the share of legacies whose parents earn above $500K according to:

$$P(\text{income} > \$500K \,|\, \text{legacy}) = \frac{P(\text{legacy} \,|\, \text{income} > \$500K) \times P(\text{income} > \$500K)}{P(\text{legacy})}.$$

See Sommeiller and Price (2018) for data on the U.S. income distribution.

[54] Remarkably, Chetty et al. (2017b) also estimate that 15.4% of students at Harvard come from families in the Top 1% of the income distribution, which is about the same number as from the bottom three quintiles combined (see p. 14). Their data come from Harvard students who graduated in the early 2000s. Chetty et al. (2017a) show that the share of Harvard students from the Top 1% has stayed steady at about 15% since 2000.

ratings.[55] Table 5 shows ratings distributions for Harvard's overall rating, as well as the profile ratings, by race and admission status, separately for non-ALDCs and LDCs. Distributions for Harvard's other ratings, such as school support measures and alumni interview scores, are given in Appendix Table D1.

The results in the top panel of Table 5 are broadly consistent with the patterns observed in the demographic characteristics. In each of the five dimensions and for each racial group, LDCs are stronger on average than the average applicant. The relative strength of LDC applicants, however, depends on the rating. The gaps are especially large for the overall, athletic, and personal ratings. On the overall rating, Hispanic LDC applicants have the lowest share of 2's at 18.5% when compared to LDC applicants of other racial groups. But this share is 3.5 times larger than the highest share for non-ALDC applicants (African Americans at 5.3%). In contrast, on the academic rating, non-ALDC Asian American applicants have a higher share of 2's than the LDC applicants of the other three races/ethnicities.

Perhaps the most striking results are for the athletic rating, which correlates strongly with LDC status and advantaged status more generally. Among LDC applicants, Asian Americans have the lowest share of 2's at 14.9%. Yet this is higher than the highest share among non-ALDC applicants (whites at 12.8%). White legacies score the highest on the athletic rating with 21.9% receiving a 2, and white non-ALDC applicants score the highest among non-ALDC applicants.

The relative strength of LDC applicants on Harvard's internal ratings is not a surprise given their family resources. What is surprising is that it is the non-academic ratings where LDC applicants particularly dominate. There are two factors working together that likely drive this pattern. First, LDC applicants are simply stronger than non-ALDC applicants in the non-academic dimensions that Harvard values. Second, when rating applicants on non-academic qualities, Harvard provides tips to LDC applicants.

While the first channel seems less controversial than the second, it is important to consider what the relative strength of LDC applicants on non-academic dimensions implies for the concept of holistic admissions. Ratings related to extracurricular activities, athletics, and

[55]For ease of exposition, we use 2 or better to refer to applicants who received any form of a 2 or 1, regardless of whether the 2 or 1 was accompanied by a plus or minus.

personal qualities are included in the admissions process to allow for a more comprehensive view of an applicant. It is widely believed that this approach—pursued by many elite US universities—opens doors for less advantaged applicants. Harvard's non-academic ratings appear to achieve that goal only if race is the central characteristic defining disadvantage. The gap between white and under-represented minority applicants is much smaller for non-academic ratings as compared to the academic rating. However, *within* racial groups, these holistic admissions criteria favor advantaged applicants. In fact, within a racial group it is generally true that the LDC and non-ALDC academic ratings are *more* similar than the non-academic ratings.

In addition to the possibility that LDC applicants have better personal qualities and overall profiles (as defined by Harvard) than non-ALDC applicants, preferences for these special applicants may also influence these non-academic ratings. According to Harvard's reader guidelines, the overall rating is essentially the reviewer's assessment of the applicant's probability of being admitted as a function of both objective and subjective factors, which according to depositions of admissions officers includes race and ALDC status (among others).[56] The reader guidelines for the personal rating provide very little guidance other than that "Outstanding" personal qualities deserve the highest rating. In fact, when OCR investigated Harvard in 1990, they found less consistency in how Harvard's readers assigned personal ratings as compared to the academic and extracurricular ratings.[57] The subjectivity inherent in the personal rating leaves greater scope for other preferences to play a role. Harvard has attempted to address this subjectivity in newly updated reader guidelines that are much more explicit about what constitutes excellence in academics, extracurriculars, and personal qualities.[58]

Somewhat surprisingly, OCR also found Harvard's athletic ratings to be quite inconsistent.[59] For example, they found that some applicants with only recreational participation

[56] See Document 419-1, pp. 172–173 and Document 421-9, pp. 253–254, 259.

[57] See Trial Exhibit P555, pp. 20–21.

[58] As mentioned in Section 2.1, Harvard altered their reader guidelines in Summer 2018 to address some of these shortcomings (see Trial Exhibit P723). In particular, readers are now instructed explicitly to ignore race when assigning the personal rating. Additionally, the description of attributes leading to the highest personal rating has increased from one word ("Outstanding") to five lines of text.

[59] See Trial Exhibit P555, pp. 20–21. Trial testimony also supports the ambiguous nature of the athletic rating. When the director of admissions was asked whether a nationally ranked figure skater would receive

during high school received a better athletic rating than applicants who participated on varsity teams in high school. This coincides with testimony from Harvard's admissions director which indicated that the athletic rating has a "team component" to it in the sense that an applicant would be rated more highly if they participated in a sport that Harvard competes in at the varsity level.[60] Applicants who play a leadership role in their sport are also likely to receive a higher athletic rating.[61] Both of these ratings advantages are likely to benefit LDC applicants since they are more likely to attend private secondary schools.[62] Relative to public secondary schools, private schools may be more likely to offer an array of varsity teams similar to Harvard's, and to provide more leadership opportunities since they tend to be smaller on average.

The comparative advantage of LDC applicants over non-ALDC applicants generally reverses when we look only at those applicants who were admitted. The bottom panel of Table 5 shows that LDC admits are weaker than non-ALDC admits in most dimensions.[63] This is not true for the athletic rating, which especially favors LDC applicants.[64]

Within racial groups, LDC applicants are on average stronger than the average applicant, but the average LDC admit is weaker than the average non-ALDC admit. This pattern suggests a significant preference for LDC applicants in the admissions process. In other words, LDC applicants within racial categories are doubly advantaged. They benefit from the particular criteria Harvard uses to evaluate applicants (see also Lee and Ries, 2019), and they receive an additional tip in admissions decisions conditional on those criteria.

a high athletic rating, she responded: "They might not get an athletic rating since we don't have a figure skating team. There are members of my staff who would think that it's not fair to indicate – for example, your daughter, were she a candidate and so on and she were still a figure skating champion, we would probably want to record that in some way in that box. Others of us would make her a 1 or a 2+ or something in the extracurricular box." See Day 6 Trial Transcript, pp. 181–183, for additional details.

[60] See Day 6 Trial Transcript, pp. 181–183.

[61] See Trial Exhibit P001. Similar to the personal rating, the reader guidelines for assigning the athletic rating have also recently been updated (see Trial Exhibit P723). The new guidelines explicitly state that a student with an opportunity to walk on to a varsity team should receive a higher rating.

[62] Klein and Keto (2015) indicate that 60% of legacy students attended a private secondary school while only 35.4% of all students attended a private secondary school.

[63] Appendix Table D2 shows the distributions of Harvard's other ratings for admitted students.

[64] There are a handful of other ratings and race combinations for which LDC admits are stronger than non-ALDC admits, but this only occurs for African Americans and Hispanics. Here it is important to note that the tip African Americans and Hispanics receive for LDC status is lower than that of non-ALDCs. We discuss this further in Section 4.1.

3.3.3 Academic Preparation

We investigate these disparities further by comparing in greater detail the academic qualifications of LDC and non-ALDC applicants. To do this, we construct deciles of Harvard's academic index, which is a weighted average of the applicant's scores on the SAT, SAT II, and high school grade point average (or class rank).[65] The deciles are defined over academic indexes for all LDC and non-ALDC applicants with non-missing values.[66] We then show in Table 6 how LDC and non-ALDC applicants are distributed across the academic index deciles by race, and report the admit rates for each racial/ethnic group and LDC status by decile.[67]

Panel A of Table 6 shows the distribution of the academic index for LDCs and non-ALDCs by race. For all racial/ethnic groups, LDC applicants are less likely to be in the bottom two deciles than their non-ALDC counterparts.[68] Moreover, for most racial/ethnic groups LDC applicants are more likely to be in the top two deciles. The exception is Asian Americans. While LDC Asian Americans have higher academic indexes than LDC applicants of other races, non-ALDC Asian American applicants have significantly higher representation in the top two deciles (34%) than their LDC counterparts (27%).

The key results are in Panel B where we show the admit rates associated with different deciles of the academic index. Consider the bottom decile. No non-ALDC white, Hispanic, or Asian Americans are admitted, and the admit rate for African Americans is 0.03%. Yet, white LDC applicants in the bottom decile are admitted at a rate of 6.32%, a rate 1.4 percentage points *higher* than the average admit rate for all white non-ALDC applicants. Indeed, white LDC applicants in the third decile from the bottom are admitted at a rate

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

[66] We exclude those who received the lowest score for converted grade point average (35). This is because converted GPAs range from 35 to 80, and there is a spike in the data at 35. It is apparent from the data that a 35 is often a result of grades being incorrectly converted. See Document 415-8 footnote 51 for details.

[67] A typical applicant in the top decile might have an SAT score of 1580 (combined math and verbal), an average SAT2 score of 790 (out of 800), and perfect grades. This combination yields a academic index value of 238. On the other hand, a typical applicant in the bottom decile might have an SAT score of 1150, an average SAT2 score of 575, and average grades (among applicants). This combination of test scores and GPA yields an academic index value of 190. See Trial Exhibit P626 for the correspondence between academic index values and deciles. Trial Exhibit DX 730 reports the average converted GPA among applicants.

[68] A surprising result is the lack of academic strength among African American LDC applicants. Nearly a quarter of this group is in the lowest academic index decile.

higher than white non-ALDC applicants in the top academic index decile.

Harvard acknowledges admissions preferences for underrepresented minorities.[69] Using this as a reference point, we see that, for the top five deciles, African American non-ALDCs have admit rates similar to white LDCs. However, in the bottom five deciles the admit rates for white LDCs are much higher. Note that the characteristics of the applicants are likely to be different across these groups even within academic index deciles. What the table shows is what the combination of preferences and observed differences in characteristics have to be in order to account for the observed differences in admit rates.

The results in Table 6 illustrate that, while LDC applicants are stronger academically than non-ALDC applicants, admit rates for less academically prepared LDC applicants are much higher. Since these patterns move in opposite directions, it is unclear how LDC admits compare to non-ALDC admits in terms of academic preparation. Table 7 fills this gap by showing how admits are distributed across the academic deciles and the share of admits within each decile that are LDC. A clear and consistent pattern emerges. The total number of admits is monotonically increasing in academic preparation; however, the share of admits that are LDC decreases dramatically as academic preparation improves.[70] For example, in the second academic index decile there are only 52 white admits, of which 67% are LDC. In contrast, there are 952 white admits in the top academic decile, of which only 20% are LDC. A similar pattern is seen across other racial groups, though it is muted, since there are so few non-white LDC applicants.

The patterns in the share of admits that are LDC across academic deciles are consistent with two ideas. First, Harvard employs large preferences for LDC applicants. Second, on non-academic measures, such as the personal, athletic, extracurricular, and school support ratings, LDC applicants are stronger. This second mechanism suggests that most of the non-academic qualifications that positively influence admission favor highly advantaged ap-

[69]Harvard's reader guidelines indicate that an applicant's overall rating is closely tied to the probability of admission and that factors that admissions officers deem valuable to Harvard, such as race, can be factored in. See Document 421-9, pp. 288, 259, 422. Additionally, the following trial documents all demonstrate substantial racial preferences for African American and Hispanic applicants: Document 415-8; Document 415-9; Document 419-141; and Document 419-143.

[70]Note that including athletes in this exercise would only exacerbate these patterns at lower academic deciles. The number of admits in the bottom deciles would be significantly higher, as would the share of admits that are ALDC.

plicants. While the standard view of holistic admissions is that it provides a channel for less advantaged applicants to obtain a leg up, within racial groups the holistic system at Harvard yields precisely the opposite.

3.4 Recruited Athlete Admits

We now turn to recruited athletes, comparing the internal Harvard ratings of admitted recruited athletes with non-ALDC applicants and admits. As discussed at the beginning of Section 3, it is more difficult to recover attributes and Harvard ratings for recruited athlete applicants from the publicly disclosed documents. For this reason, we focus on the profile ratings for admitted recruited athletes. This is not particularly limiting, since the admit rate for recruited athletes is above 85%.

Table 8 compares the ratings received by recruited athlete admits to those received by non-ALDC applicants and admits, focusing on the overall rating and the four profile ratings with the exception of the athletic rating.[71] With the publicly available documents, we are able to form meaningful bounds on the ratings of recruited athletes.[72]

Recruited athlete admits are universally weaker than non-ALDC admits on these ratings. This is not surprising, given that we know athletes are stronger on the athletic rating. But for some race and rating combinations, the differences are striking. At most, 28% of white athlete admits receive a 2 or higher on the academic rating. In contrast, 89% of white non-ALDC admits receive a 2 or higher on the academic rating. 78% of Asian American non-ALDC admits receive a 2 or higher on the extracurricular rating, while at most 12% of admitted Asian American athletes receive a similarly high extracurricular rating.

In many cases—and in contrast to LDC admits—recruited athlete *admits* are weaker than non-ALDC *applicants*. Non-ALDC applicants on average have higher academic and extracurricular ratings than admitted recruited athletes of the same race for all groups except Hispanics on the academic rating. On the overall and personal ratings, however, recruited athlete admits do have higher scores than non-ALDC applicants.

[71] The athletic rating is by definition 1 for recruited athletes.

[72] Additional details are provided in Appendix B. For non-ALDC applicants and admits, no range is necessary as these numbers are exact. Comparisons to recruited athlete applicants, as well as comparisons on the other Harvard ratings, are available in Appendix Table D3.

This last result is consistent with the patterns observed for LDC applicants. LDC applicants are much stronger than non-ALDC applicants precisely in the overall, personal, and athletic ratings. We know that LDC and recruited athlete applicants have similar racial distributions, but this suggests that recruited athletes may be more similar to LDC applicants than non-ALDC applicants in additional dimensions. We investigate this further by examining the share of admitted athletes that are disadvantaged. We find that at most 3.2% of white admitted athletes are economically disadvantaged.[73] For non-ALDC admits, the corresponding number is 14.6%.[74] The advantaged status of recruited athletes can be verified using *The Harvard Crimson* survey cited earlier.[75] For the Class of 2019, 26% of recruited athlete respondents came from families earning more than $500,000 in income per year. The corresponding numbers for legacies and all respondents are 40.7% and 15.4%, respectively.

ALDC applicants, including athletes, are highly advantaged relative to non-ALDC applicants and are admitted at much higher rates. ALDC admits are also significantly weaker than non-ALDC admits on almost all of Harvard's internal ratings. These patterns strongly suggest that Harvard employs significant preferences for these advantaged groups when making admissions decisions. However, the descriptive work thus far cannot rule out the possibility that recruited athletes and LDC applicants are stronger once we account for all characteristics jointly and in the manner that Harvard weighs them in the admissions decision. In the following sections we investigate this possibility by estimating a model of Harvard admissions.

4 Estimates of the Legacy and Athlete Advantage

To better understand the source of the gap in admission rates between ALDC and non-ALDC applicants, we estimate a logistic regression model of Harvard's admissions decisions. A full discussion of the admissions model we estimate, as well as a list of all controls, is given in

[73] This number can be derived using Table B.3.2 from Document 415-8 and Table B.3.2R from Document 415-9.

[74] A similar pattern exists for the other racial groups—the upper bound for athletes is well below the value for non-ALDC admits. We focus on white athletes since they account for close to 70% of admitted recruited athletes (see Panel C of Table 3).

[75] See Klein and Keto (2015).

Appendix C.[76] Here, we provide a brief overview.

The Harvard admissions data cover six admissions cycles and include hundreds of variables describing each applicant. It is not feasible to include every variable separately in every year, as we would ultimately have as many regressors as admits. We choose to pool the application cycles together and estimate a single logistic regression with indicators for each admissions cycle. Estimating the admissions model with indicators for each admissions cycle ensures that the predicted number of admits matches the actual number of admits.

Including only these admissions cycle indicators to capture the time-varying component of Harvard's admission process would be consistent with Harvard having a single index of applicant quality in every year. Yet, there is evidence that Harvard cares about the admitted composition of the class. 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.[77] Admissions officers can use these "one-pagers" to generate similarly constituted admit classes over time, even if the applicant pool is changing.[78] To capture these effects, we include interactions of many of the variables on the one-pagers with admission cycle indicators.

We incorporate a broad set of applicant controls in the model, including numerous measures of socioeconomic status, neighborhood and high school attributes, region, intended major, and academic aptitude, among others. We incorporate 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 not included because the rating is specifically designed to incorporate admissions preferences; the personal rating is not included because there is empirical evidence that is influenced by preferences as well.[79]

[76] Additional details are provided in Section 3.7 of Document 415-8 and Section 8 of Document 415-9.

[77] See Trial Exhibit P164 for the characteristics that Harvard tracks as they make admissions decisions.

[78] Arcidiacono, Kinsler, and Ransom (2019) show that the fraction of admitted students that are legacy and athlete shows no time trend over an 18-year period, despite legacies and athletes making up a decreasing share of applicants over time.

[79] For example, Document 415-9 shows that, in contrast to the other ratings (with the exception of the overall rating), an ordered logit model of the personal rating shows that legacies receive a bonus for this rating (see Table B.6.7R).

To allow for the possibility that racial preferences operate differently for LDC applicants, we also interact each of the LDC categories with race. For similar reasons, we interact race with a number of other characteristics, such as disadvantaged status and gender.[80]

Our preferred model excludes athletes. With the admit rate of athletes at 86% compared to less than 5.5% for non-ALDCs, what variables matter and how they matter is likely to be different for athletes.[81] In total, our preferred model controls for 370 variables and yields a Pseudo R^2 of 0.56. By comparison, the models in Espenshade, Chung, and Walling (2004) achieve a Pseudo R^2 no larger than 0.20.[82]

Appendix Table D4 displays a subset of the estimated coefficients as we vary the set of controls in the model. Model (5) is our preferred model. Since we interact variables such as legacy and dean's interest with year, all coefficients are for the base year, the Class of 2014.[83]

4.1 Transformations

The estimated coefficients on indicators for legacy, double legacy (i.e. both parents are alumni), faculty or staff child, and being on the dean's interest list are all large, positive, and statistically significant. To put the size of the LDC preferences in context, we examine how the probability of admission would change for non-ALDC applicants of different races and different baseline probabilities of admission using the results from our preferred admissions

[80] Arcidiacono (2005) shows that racial preferences for African Americans in admissions and financial aid vary with whether the applicant is low income. 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. See Document 415-9, Table B.3.2R.

[81] As we show in section 5, this is indeed the case. Factors such as the academic rating and extracurricular rating become less important when athletes are included in the model.

[82] The Pseudo R-squared is a different measure from the R-squared, and the two are not interchangeable. As measures, the only similarities the two have is that they run on a scale from 0 to 1, and higher values on either are indicative of a better fit of the data. The classic citation for what is considered an "excellent fit" based on the Pseudo R^2 is McFadden (1979), p. 307:

> Those unfamiliar with the ρ^2 index should be forewarned that its values tend to be considerably lower than those of the R2 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 above later became known as McFadden's R-Squared, or the Pseudo R-squared.

[83] While the coefficients on legacy interacted with year were included in the publicly released reports, other interactions between year and special recruiting categories were not reported. Hence, we focus on the Class of 2014 for the transformation exercises in the next subsection.

model.[84] Consider, for example, how the probability of a white non-ALDC applicant with a baseline probability of admission of X would change if they happened to be a legacy, but if all other characteristics remained the same. If the baseline probability of admission is X, we can calculate the index of observables, Z, for this applicant according to the log odds formula

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

which is the inverse of the standard logit formula. If the applicant were a legacy, we simply add the coefficient on the legacy indicator (2.141) so that the new admissions index would be $Z + 2.141$. The new admissions probability would then be given by $\frac{\exp(Z+2.141)}{1+\exp(Z+2.141)}$. A similar calculation can be done for other racial groups, where we also need to account for the race-legacy interaction coefficients.

Table 9 examines how the probability of admission would change for various hypothetical non-ALDC applicants. A white applicant with an initial admit probability of 5% would see their admissions probability increase to 31% if they were a legacy. This number would increase to 47% if they were a double legacy. While these preferences are large, the preferences for those on the dean's interest list are even larger at 59%. The final column shows that these preferences are much larger than those given to disadvantaged students, where the corresponding probability is 21%.

African Americans also see increases in their probability of admission through being treated as a legacy or on the dean's interest list. However, the gains are muted. An African American non-ALDC applicant with an initial admit probability of 5% would see their admissions probability increase to 16% if they were a legacy and to 29% if they were a double legacy. African Americans do not receive the full legacy tip. This may be because they already receive a large tip for being African American. Indeed, the base African American tip is slightly larger than the preference given to those on the dean's interest list. This same feature is also present in the last column, which illustrates that disadvantaged African Americans receive virtually no tip for being disadvantaged.

[84] Our preferred model excludes the personal rating since there is evidence that the personal rating incorporates preferences for ALDC applicants. Using a model that includes the personal rating does not alter the basic findings.

The previous results are based on an admissions model that excludes athletes. However, we also estimate a slightly altered admissions model that includes athletes. This alternative model contains fewer controls, but is otherwise very similar to the model outlined in the previous section.[85] Appendix Table D5 lists the coefficients from the model including athletes. The coefficient on the recruited athlete variable is substantially larger than any of the other previously discussed preferences, regardless of the controls we include. To make this more precise, consider a white, non-ALDC applicant who has only a 1% chance of admission. If this applicant were treated as a recruited athlete, the admission probability would increase to 98%. Being a recruited athlete essentially guarantees admission even for the least-qualified applicants. A similar calculation, but in reverse, emphasizes the advantage athletes receive. An athlete who has an 86% probability of admission—the average rate among athletes—would have only a 0.1% chance of admission absent the athlete tip.

4.2 Removing Preferences

The transformation exercises in the previous section tell us how the admit probability for a hypothetical applicant would change with legacy or athlete status. But because we do not know the full distribution of attributes for ALDC applicants (nor do we report all coefficients in the model), we cannot determine what would happen to the admissions chances of ALDC applicants as a whole if those preferences were removed. However, calculations presented in Exhibit 287 allow us to address this question. The admissions model used in this public document is quite similar to the model used to calculate LDC preferences. The one difference is that athletes were added back into the model and athlete was interacted with race.

The first sets of rows of Table 10 show the total number of admits and applicants overall and by LDC and athlete status. These rows are followed by admission rates by ALDC status. The non-ALDC admit rates range from 4.9% (white) to 7.6% (African Americans). For ALDC applicants, the admit rates range from 42.6% (Hispanics) to 46.7% (African Americans). The exhibit makes it possible to show how much of this admit rate gap between ALDCs and the overall population is the result of ALDC preferences.

[85] See Figure 7.1 in Document 415-8 for a detailed description of the controls in the model. The relevant sample is discussed in Section 2.3.3.

Namely, Exhibit 287 shows what the admit rate would be for previous admits if ALDC preferences were removed. The exhibit reports the share of admits, by race, that would still be admitted if ALDC preferences were removed. This calculation follows directly from Bayes' rule. Denote $y = 1$ if an applicant was admitted when a preference was in place. Denote $y' = 1$ as an indicator for whether an applicant would be admitted when the preference is removed. The probability an applicant would still be admitted after the preference is removed can be written as:

$$\Pr(y' = 1 \,|\, y = 1) = \frac{\Pr(y = 1 \,|\, y' = 1)\Pr(y' = 1)}{\Pr(y = 1)}. \tag{2}$$

The first term on the right hand side is, by definition, 1: if an applicant was admitted without a preference, the applicant will also be admitted when a preference is in place. The other two terms are the model-predicted probabilities without and with the preference.

The bottom part of Table 10 shows the results of this exercise. Removing ALDC preferences would result in an admit rate of 68% for white applicants previously admitted when ALDC preferences were in place. Note that, by definition, this drop occurs solely for those in the ALDC category. For whites, the share of applicants who are ALDC is 8%. With the drop in the admit rate occurring only for this group, the share of white ALDC admits who would be admitted if they were instead treated as other white applicants would be 26%.

Table 10 also shows what this implies for the overall admit rate for ALDC applicants. The admit rate for all white ALDCs applicants would fall from 43.6% to 11.4%, a drop of over 30 percentage points. We can go one step further and generate an upper bound on the overall LDC admit rate for white applicants when ALDC preferences are eliminated. If we assume that all of the athletes are rejected when ALDC preferences are eliminated, it would imply that the white LDC admit rate would fall from 33.4% to less than 14%. Thus, the average marginal effect of being an LDC on admissions for white applicants is at least 19.4%.

5 Counterfactual Simulations

While the transformation examples and the exercises utilizing Bayes' rule in the previous section highlight the importance of ALDC preferences, neither approach accounts for capacity constraints. For example, if ALDC preferences are eliminated, the likelihood of ALDC applicants being admitted would decline, and, with no other change, the class size at Harvard would shrink. However, using the admissions model, we can correct for these capacity concerns.[86] In this section, we describe our approach and investigate how the racial distribution of the admitted class would change if these preferences were removed. In describing the methodology, we focus on legacy preferences for ease of exposition.

5.1 Methodology

We use the estimated admission model to calculate an admission probability for each applicant in the absence of legacy preferences by setting the coefficients related to legacy status to zero and keeping all other coefficients the same. This rules out Harvard changing the weights on characteristics that are correlated with legacy status to undo the removal of legacy preferences. We then aggregate these probabilities up to construct an expected class. In order for the estimated model to match the racial distribution of the admitted class in every admission cycle, the estimated admissions model included race-by-year interactions. The race-by-year coefficients will ensure that the estimated model perfectly matches the actual number of admits in each racial group in every year.

One approach to deal with capacity concerns would be to rank applicants according to their admissions index sans legacy preferences, and then select the highest ranked applicants such that the number of admits matches the observed data in every cycle. However, this would treat admissions decisions as if they were deterministic based on the admissions index, which is incompatible with the assumptions of the logit model and with what we know about how Harvard evaluates applications.[87] Instead, we adjust the admissions index of

[86] Document 415-9 shows this for legacy preferences and athlete preferences separately, as well as the combination of removing legacy, athlete, and racial preferences.

[87] Said differently, the admissions index summarizes one's admissions likelihood based on observable characteristics, but we know that unobservable factors also influence the admissions outcome.

all applicants by the same amount such that the average admission probability when legacy preferences are eliminated matches the average admission probability with legacy preferences. Numerically, this means that we solve for an index adjustment ϕ_t^* in each admissions cycle t, such that

$$\overline{p}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\exp(X_i \hat{\beta}_{NL} + \phi_t^*)}{1 + \exp(X_i \hat{\beta}_{NL} + \phi_t^*)} \tag{3}$$

where $\overline{p}_t$ is the actual average probability of admission in admission cycle t, N_t is the size of the relevant applicant pool in cycle t, X_i reflects the characteristics of applicant i, and $\hat{\beta}_{NL}$ are the estimated coefficients on these characteristics with the coefficients on legacy and all legacy interactions set to zero. Finding the ϕ_t^* that solves this equation guarantees that when we aggregate the individual admissions probabilities under the assumption of no legacy preferences, we maintain the exact number of admits each year. The composition of the class will change, however, since different racial groups will experience heterogeneous changes to their admissions probabilities.[88]

Even though athletes are not part of our preferred model, we can still use the model to evaluate the removal of athlete or legacy preferences. In the case of removing legacy preferences, we simply treat the decisions for athletes as fixed. In the case of removing athlete preferences, we treat athletes like any other applicant and calculate a predicted admission probability from the admissions model.[89]

5.2 Counterfactual Results

The results of this exercise split out by applicant race are displayed in Table 11. The first row for each racial group shows the model-generated number of admits aggregated across the six years. The model predictions match the racial composition of admits in the data, since the estimated admissions model includes race-by-year interactions. The second row illustrates how the number of admits changes when preferences for legacy applicants are eliminated. The only group of applicants that experiences a decline in the number of admits is white

[88] We do not model the equilibrium impact on the application margin, because our data only includes applicants to Harvard and does not include applicants who did not apply to Harvard.

[89] When we eliminate athlete preferences, we change both the athletic and extracurricular ratings to 2 for recruited athletes. See page 9 of Appendix A in Document 415-9 for additional details.

applicants. The number of white admits falls by approximately 4%, while the number of African American, Hispanic, and Asian American admits respectively increase by 4%, 5%, and 4%. The third row illustrates how the racial composition of the class changes when athlete preferences are eliminated. In this case, the number of white admits declines by 6%, while the number of Hispanic and Asian American admits rises by 7% and 9% respectively. African American admits are essentially unchanged.[90]

To be clear, this does not imply that all whites are hurt by the removal of legacy and athlete preferences, nor do all Asian Americans benefit. The aggregate changes in white enrollments mask within-race shifts from legacy and athlete admits to those who are not legacies and athletes.

When legacy or athlete preferences are eliminated, we estimate that the racial composition of Harvard's admitted class changes by a non-trivial amount. Yet, we believe that the numbers reported above likely understate the true impact of eliminating legacy and athlete preferences on the racial composition of the admitted class. First, while we are able to report counterfactuals for the elimination of legacy or athlete preferences, we cannot report what would happen if both of these preferences were eliminated. Moreover, we cannot run the counterfactual when all ALDC preferences are eliminated. We suspect that if we were able to run these counterfactuals, the share of white admits would drop by significantly more than 6% and the share of Asian American admits would rise by more than 9%.

Second, when we estimate our admissions model, we include LDC applicants as part of the estimation sample. However, the way characteristics for LDC applicants matter for admissions may be different than how those same characteristics matter for non-ALDC applicants. For example, Table 6 shows that, for the bottom decile of the academic index, *no* white, Hispanic, or Asian American applicants were admitted. Yet, white LDC applicants in this decile were admitted at a higher rate (6.35%) than the average across all non-ALDC applicants (5.46%).

[90]The last row in Table 11 shows what would happen if in addition to removing legacy and athlete preferences, we also removed racial preferences. In this case, the coefficients on legacy, athlete, and race/ethnicity are set to zero as well as their interactions. The counterfactual shows that the number of admitted African Americans would be a third of what it was when all these preferences were in place. The number of admitted Hispanics would decline by almost half. Clearly the preferences African Americans and Hispanics receive do not simply offset the losses they incur from legacy and athlete preferences.

This phenomenon is illustrated in Table 12, where we show how the coefficients on the academic rating and the extracurricular rating change when athletes and then LDC applicants are excluded from the model.[91] The boost an applicant receives from obtaining an academic rating of 1 relative to an academic rating of 3 is higher in a model with only non-ALDC applicants as compared to a model that also includes LDC applicants. The importance of an academic rating of 1 declines even further if athletes are included. The change in the coefficients is even more dramatic at the bottom of the academic rating. The penalty for receiving a 4 on the academic rating is much smaller when the model is estimated with LDC applicants included than when estimated only on non-ALDC applicants and becomes smaller still if recruited athletes are included as well. Similar patterns hold for the extracurricular rating, with the coefficients becoming attenuated when LDC applicants and recruited athletes are included in the model.[92]

In our counterfactuals, we avoided the athlete-generated coefficient distortions by excluding them from the admissions model. We were are still able to evaluate the impact of removing athlete preferences by treating their admissions decisions as fixed except when we eliminate athlete preferences, in which case we used their characteristics and the coefficients of the model to predict their admissions probabilities. Ideally, we would have pursued a similar strategy for the LDC groups. We suspect that the changes in the racial composition would have been even more stark in this case.

6 Conclusion

Detailed data on how universities practice holistic admissions are virtually never made available to researchers. Through the SFFA lawsuit, unprecedented access was given to how Harvard rates their applicants as well as how applicant characteristics, including these ratings, translate into admissions.

[91] The source for this table is Table 2 of Exhibit 287.

[92] As mentioned by Norton and Dowd (2018), it is not possible to make direct comparisons of logit coefficients across specifications or subsamples, because the coefficients depend on the variance of unobservables (which depends on the model specification and data subsample). When we account for this property of logit models by dividing each coefficient by the Academic Rating=1 coefficient, the results in Table 12 are nearly identical. This is because the variance of unobservables is not sensitive to the inclusion of LDCs or athletes in a model as rich as ours.

This paper has focused on the substantial preferences ALDC applicants receive. The advantages for athletes are especially large, with an average admit rate for recruited athletes of 86%. This high admit rate occurs despite admitted athletes often being worse on Harvard's ratings than the applicant pool itself. Overall, our results show that only one quarter of white ALDC admits would have been admitted if they had been treated as a non-ALDC applicant.

Each of the ALDC preferences primarily benefit white students. Over 43% of white admits are ALDC, compared to less than 16% of admits for each of the other three major racial/ethnic groups. Indeed, due in part to the nature of the sports that Harvard offers, recruited athletes alone make up over 16% of white admits. We show that removing legacy and athlete preferences results in shifts in admissions away from white applicants with each of the other groups either increasing or staying the same. At the same time, fewer high-income applicants would be admitted.

Harvard—and other institutions that use holistic admissions criteria—may benefit from employing ALDC preferences, both through donations and through providing amenities to its student body. But given that the beneficiaries of these practices already come from quite advantaged backgrounds, and the further evidence that these preferences appear to be increasing over time (Arcidiacono, Kinsler, and Ransom, 2019), the exposure of the scope of these practices may lead them to be reevaluated.

References

Anderson, Michael L. 2017. "The Benefits of College Athletic Success: An Application of the Propensity Score Design." *Review of Economics and Statistics* 99 (1):119–134.

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

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

Bound, John, Brad Hershbein, and Bridget Terry Long. 2009. "Playing the Admissions Game: Student Reactions to Increasing College Competition." *Journal of Economic Perspectives* 23 (4):119–146.

Chappell, Bill and Merrit Kennedy. 2019. "U.S. Charges Dozens Of Parents, Coaches In Massive College Admissions Scandal." *NPR* URL https://www.npr.org/2019/03/12/702539140/u-s-accuses-actresses-others-of-fraud-in-wide-college-admissions-scandal. Accessed March 23, 2019.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2017a. "Mobility Report Card: Harvard University." Working paper, UC Berkeley. URL http://eml.berkeley.edu/~yagan/college/Harvard_University.pdf.

———. 2017b. "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility." Working Paper 23618, National Bureau of Economic Research.

Clotfelter, Charles T. 2003. "Alumni Giving to Elite Private Colleges and Universities." *Economics of Education Review* 22 (2):109–120.

Dale, Stacy Berg and Alan B. Krueger. 2014. "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data." *Journal of Human Resources* 42 (2):323–358.

Day 3 Trial Transcript. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialTranscripts/20181017-Day-3.pdf.

Day 5 Trial Transcript. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialTranscripts/20181019-Day-5.pdf.

Day 6 Trial Transcript. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialTranscripts/20181022-Day-6.pdf.

DeSilver, Drew. 2019. "A Majority of U.S. Colleges Admit Most Students Who Apply." Report, Pew Research Center. URL https://www.pewresearch.org/fact-tank/2019/04/09/a-majority-of-u-s-colleges-admit-most-students-who-apply/.

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.

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.

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.

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.

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.

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.

Document 611. 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://cases.justia.com/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/611/0.pdf?ts=1541565114. List of exhibits that were admitted into evidence during trial.

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.

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.

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

Jackson, Abby. 2017. "Former Ivy League Admissions Officer Reveals How Schools Pick Students." *Yahoo! Finance* URL https://finance.yahoo.com/news/former-ivy-league-admissions-officer-150000242.html. Accessed June 24, 2019.

Jacob, Brian, Brian McCall, and Kevin Stange. 2018. "College as Country Club: Do Colleges Cater to Students' Preferences for Consumption?" *Journal of Labor Economics* 36 (2):309–348.

Klein, Mariel A. and Elizabeth C. Keto. 2015. "Class of 2019 by the Numbers." *The Harvard Crimson* URL https://features.thecrimson.com/2015/freshman-survey/makeup-narrative/. Accessed April 9, 2019.

Lee, Sanghoon and John Ries. 2019. "Noncognitive Skills in Admission Decisions: Impacts on Demographics, Academic Performance, and Labour Market Outcomes." Working paper, University of British Columbia.

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.

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

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

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

Norton, Edward C. and Bryan E. Dowd. 2018. "Log Odds and the Interpretation of Logit Models." *Health Services Research* 53 (2):859–878.

Riehl, Evan. 2018. "Fairness in College Admission Exams: From Test Score Gaps to Earnings Inequality." Working paper, Cornell University.

Sabky, Rebecca. 2017. "Check This Box if You're a Good Person." *The New York Times* URL https://www.nytimes.com/2017/04/04/opinion/check-this-box-if-youre-a-good-person.html. Accessed June 27, 2019.

Sommeiller, Estelle and Mark Price. 2018. "The New Gilded Age: Income Inequality in the U.S. by State, Metropolitan Area, and County." Report, Economic Policy Institute. URL https://www.epi.org/publication/the-new-gilded-age-income-inequality-in-the-u-s-by-state-metropolitan-area-and-county

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.

Trial Exhibit DX 005. 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/D005.pdf. Alumni Interviewer Handbook.

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.

Trial Exhibit DX 706. 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/D706.pdf. Share of admitted students in ALDC categories.

Trial Exhibit DX 730. 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/D730.pdf. Academic qualifications and profile ratings of transfer applicants.

Trial Exhibit DX 746. 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/D746.pdf. Racial composition of admitted ALDC students.

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.

Trial Exhibit P104. 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/P104.pdf. Email exchange discussing admission of donors' children.

Trial Exhibit P106. 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/P106.pdf. Email exchange rating a donor as a "2".

Trial Exhibit P164. 2018. In *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College et al.* Civil Action No. 14-14176-ADB (D. Mass). URL https://github.

com/tyleransom/SFFAvHarvard-Docs/blob/master/TrialExhibits/P164.pdf. One-pager.

Trial Exhibit P316. 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/P316.pdf. Report of the Committee to Study Race-Neutral Alternatives.

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.

Trial Exhibit P618. 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/P618.pdf. Admit rates for ALDCs and non-ALDCs by academic rating.

Trial Exhibit P619. 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/P619.pdf. List of applicants interviewed and admitted.

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.

Trial Exhibit P623. 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/P623.pdf. Ratings frequencies for expanded sample.

Trial Exhibit P626. 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/P626.pdf. Number and share of applicants by academic index decile.

Trial Exhibit P723. 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/P723.pdf. Harvard Reading Procedures for Class of 2023.

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

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

Figures and Tables

Table 1: Number of Applicants and Admit Rate by ALDC Status

Group	Number of Applicants	Admit Rate (%)
Not Athlete	165,353	6.0
Athlete	1,374	86.0
Not Legacy	162,083	5.9
Legacy	4,644	33.6
Not Dean's Interest List	164,226	6.1
Dean's Interest List	2,501	42.2
Not Child of Faculty or Staff	166,406	6.6
Child of Faculty or Staff	321	46.7

Source: Data presented in Table A.2 of Document 415-8.

Table 2: Admit Rates of Harvard Applicants by ALDC Status and Academic Rating

Academic Rating		Non-ALDC	LDC	Athlete
1	# of Applicants	612	60	1
	# of Admits	405	58	1
	Admit Rate (%)	66.18	96.67	100.00
2	# of Applicants	59,731	3,118	303
	# of Admits	5,986	1,528	291
	Admit Rate (%)	10.02	49.01	96.04
3	# of Applicants	57,874	2,444	821
	# of Admits	1,390	442	716
	Admit Rate (%)	2.40	18.09	87.21
4	# of Applicants	18,176	373	210
	# of Admits	3	13	167
	Admit Rate (%)	0.02	3.49	79.52
5	# of Applicants	6,335	46	8
	# of Admits	0	0	4
	Admit Rate (%)	0.00	0.00	50.00
Total	# of Applicants	142,728	6,041	1,343
	# of Admits	7,784	2,041	1,179
	Admit Rate (%)	5.45	33.79	87.79

Source: Data presented in Trial Exhibit P618. The final set of rows is computed by the authors using the information in the preceding sets of rows.

Table 3: Admissions Statistics by Race and ALDC Status

	Non-ALDC	Athlete	Legacy	Dean's List	Faculty/Staff
Panel A: Admission Rates					
White	4.89	87.94	34.07	41.96	45.78
African American	7.58	86.11	28.57	32.53	20.00
Hispanic	6.16	88.52	35.63	41.79	42.11
Asian American	5.13	87.07	35.14	47.83	47.56
Panel B: Racial Distribution of Applicants by ALDC Status					
White	40.34	69.28	68.66	68.29	53.21
African American	10.97	10.74	5.28	3.58	3.21
Hispanic	12.59	4.55	5.65	5.77	6.09
Asian American	28.32	8.65	10.54	11.89	26.28
Panel C: Racial Distribution of Admits by ALDC Status					
White	36.15	69.30	69.17	67.17	52.41
African American	15.25	10.52	4.46	2.73	1.38
Hispanic	14.22	4.58	5.95	5.66	5.52
Asian American	26.62	8.57	10.95	13.33	26.90
Panel D: Proportion of Total Admits by Race					
White	56.36	16.36	20.49	13.32	1.52
African American	85.27	8.91	4.74	1.94	0.14
Hispanic	86.28	4.21	6.86	4.36	0.62
Asian American	84.81	4.13	6.63	5.40	1.60

Notes: All numbers in this table are percentages. Panels B and C should be read vertically, while Panel D should be read horizontally. The columns in Panels B and C add to less than 100% since there are other racial groups not shown.
Source: Authors' calculations from data presented in Tables B.3.1R and B.3.2R of Document 415-9. See Appendix B.2.3 for a complete discussion of the calculations.

Table 4: Applicant Characteristics by Race and LDC Status

	White		African American		Hispanic		Asian American		Total	
	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC
Panel A: All Applicants										
Admitted	4.89	33.42	7.58	27.46	6.16	34.66	5.13	37.14	5.45	33.79
Disadvantaged	6.36	0.62	29.21	7.78	24.33	4.81	10.85	2.37	12.58	1.79
First-generation college	4.28	1.13	13.90	3.48	21.90	3.02	8.07	1.71	8.76	1.37
Applied for fee waiver	8.20	1.54	42.63	13.57	35.58	6.33	13.16	3.03	17.08	2.47
Financial Aid	73.75	35.25	93.46	63.25	88.56	42.45	76.41	35.36	78.45	37.39
Application read by 3rd reader	15.03	62.65	18.02	50.47	18.90	63.06	16.48	63.23	16.41	62.08
N	57,582	4,075	15,664	295	17,970	352	40,415	727	142,728	6,041
Share of applicants who are LDC	6.61		1.85		1.92		1.77		4.06	
Panel B: Admitted Applicants										
Disadvantaged	14.61	0.57	28.48	3.59	37.40	3.25	21.86	3.38	22.11	1.46
First-generation college	4.05	1.17	7.67	0.00	19.96	0.82	9.65	1.15	8.54	1.17
Applied for fee waiver	12.15	0.90	28.14	6.22	35.59	1.64	18.39	2.17	19.75	1.36
Financial Aid	72.17	24.25	90.73	48.15	88.98	27.03	77.27	27.74	79.07	25.73
Application read by 3rd reader	94.78	90.52	95.11	84.00	97.11	90.16	96.14	89.29	95.57	89.75
N	2,814	1,362	1,187	81	1,107	122	2,072	270	7,784	2,041
Share of admits who are LDC	32.61		6.39		9.93		11.53		20.77	

Note: All variables are binary and the reported means are expressed as percentages.
Source: Authors' calculations from data presented in Tables B.3.1R and B.3.2R of Document 415-9. See Appendix B.2.4 for a complete discussion of the calculations.

Table 5: Share of Applicants and Admits Receiving a 2 or Higher on Various Ratings, by Race and LDC Status

	White		African American		Hispanic		Asian American	
	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC
Panel A: Applicants								
Overall	4.43	21.60	5.29	19.66	3.88	18.47	4.84	25.58
Academic	45.29	54.43	9.19	15.25	16.74	41.19	60.21	63.27
Extracurricular	24.35	36.22	15.54	30.85	16.83	31.53	28.23	37.83
Athletic	12.79	21.89	6.82	15.93	7.51	18.18	4.81	14.86
Personal	21.27	40.88	19.01	40.68	18.68	38.92	17.64	35.49
Panel B: Admits								
Overall	59.70	57.27	59.14	62.96	50.14	50.00	62.36	62.22
Academic	88.77	78.34	59.39	43.21	65.40	70.49	94.40	85.56
Extracurricular	73.03	55.80	51.98	55.56	56.64	50.00	78.28	60.37
Athletic	20.97	27.90	14.24	28.40	15.27	23.77	7.19	18.15
Personal	83.76	70.19	74.39	80.25	77.87	63.93	73.26	60.37

Source: Authors' calculations from data presented in Trial Exhibit P621 and Trial Exhibit P623. See Appendix B.2.5 for a complete discussion of the calculations.

Table 6: Shares and Admission Rates of Applicants by LDC Status, Race, and Academic Index Decile

	White		African American		Hispanic		Asian American		Total	
	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC
Panel A: Share of Applicants in each Decile										
1	4.91	3.10	37.95	25.76	19.98	4.83	3.75	2.34	10.25	4.21
2	7.67	7.03	23.08	20.00	20.94	12.50	5.07	5.78	10.30	7.82
3	10.57	9.88	14.68	14.92	16.32	13.64	6.56	7.15	10.55	10.24
4	11.07	11.62	8.24	10.17	12.17	13.07	7.49	9.22	9.74	11.45
5	13.33	13.22	5.75	15.93	9.59	12.78	9.61	12.93	10.84	13.22
6	10.31	12.21	3.26	3.73	6.01	11.65	8.97	11.55	8.51	11.73
7	12.28	12.36	2.85	5.76	5.29	9.94	11.23	11.83	9.94	11.56
8	11.28	11.35	2.09	1.36	4.57	8.52	13.19	12.38	10.01	10.57
9	9.95	10.88	1.26	0.34	3.01	7.10	16.21	13.62	10.05	10.54
10	8.64	8.35	0.85	2.03	2.12	5.97	17.92	13.20	9.83	8.66
Total	57,451	4,070	15,601	295	17,930	352	40,308	727	142,356	6,036
Panel B: Admit Rates of Applicants in each Decile (%)										
1	0.00	6.32	0.03	3.19	0.00	0.00	0.00	6.29	0.01	5.27
2	0.39	12.20	1.03	6.61	0.32	11.54	0.20	7.16	0.53	10.47
3	0.56	16.67	5.19	25.36	1.95	8.15	0.64	11.53	1.65	15.56
4	1.82	22.62	12.76	39.94	5.50	30.20	0.86	23.90	3.29	23.72
5	2.57	26.18	22.41	48.92	9.13	42.45	1.86	21.28	4.40	28.39
6	4.20	31.85	29.72	54.73	13.65	41.46	2.49	29.78	5.64	33.70
7	4.79	36.04	41.12	82.43	17.28	48.49	3.98	40.45	6.61	38.51
8	7.53	47.49	44.48	75.01	22.93	49.85	5.12	53.17	8.22	47.66
9	10.77	56.94	54.59	99.90	26.16	43.98	7.55	56.45	10.40	56.67
10	15.27	57.07	56.06	83.43	31.32	95.10	12.69	63.02	14.58	60.64
Total	4.90	33.47	7.58	27.52	6.16	34.73	5.14	36.75	5.46	33.73

Source: Authors' calculations from data presented in Tables 5.1R, 5.2R, B.5.1R and B.5.2R of Document 415-9. See Appendix B.2.6 for a complete discussion of the calculations.

Table 7: LDC Share of Admits by Academic Index Decile and Race

	White		African American		Hispanic		Asian American		Total	
	% LDC	Admits	% LDC	Admits	% LDC	Admits	% LDC	Admits	% LDC	Admits
1	100.00	8	57.69	4	—	0	100.00	1	90.17	15
2	67.01	52	9.52	41	29.71	17	42.36	7	38.87	127
3	66.33	101	8.58	130	6.41	61	26.16	23	27.97	344
4	48.04	223	6.81	176	10.38	134	38.14	42	26.44	620
5	41.72	338	10.26	224	10.85	176	21.73	92	25.03	905
6	38.88	407	3.83	157	10.37	164	21.75	115	25.89	922
7	34.92	519	7.11	197	9.38	181	16.18	215	22.33	1204
8	31.02	707	2.03	148	7.37	203	14.95	320	20.61	1476
9	29.06	868	0.93	108	7.23	152	10.18	549	19.50	1848
10	20.38	952	6.34	79	14.37	139	6.19	977	13.46	2357

Source: Authors' calculations from data presented in Tables 5.1R, 5.2R, B.5.1R and B.5.2R of Document 415-9. See Appendix B.2.7 for a complete discussion of the calculations.

Table 8: Upper and Lower Bounds on the Share of Applicants and Admits Receiving a 2 or Higher on Various Ratings, by Athlete Status and Race

	White			African American			Hispanic			Asian American		
	Non-ALDC		Recruited Athlete	Non-ALDC		Recruited Athlete	Non-ALDC		Recruited Athlete	Non-ALDC		Recruited Athlete
Rating	Applicant	Admit	Admit	Applicant	Admit	Admit	Applicant	Admit	Admit	Applicant	Admit	Admit
Overall												
Lower bound	0.0443	0.5970	0.2705	0.0529	0.5914	0.1290	0.0388	0.5014	0.1481	0.0484	0.6236	0.2376
Upper bound			0.3035			0.1935			0.3333			0.3960
Academic												
Lower bound	0.4529	0.8877	0.2460	0.0919	0.5939	0.0000	0.1674	0.6540	0.1111	0.6021	0.9440	0.4158
Upper bound			0.2791			0.0645			0.2963			0.5743
Extracurricular												
Lower bound	0.2435	0.7303	0.0832	0.1554	0.5198	0.0645	0.1683	0.5664	0.0000	0.2823	0.7828	0.0000
Upper bound			0.1163			0.1290			0.1667			0.1188
Personal												
Lower bound	0.2127	0.8376	0.4676	0.1901	0.7439	0.6290	0.1868	0.7787	0.5185	0.1764	0.7326	0.2871
Upper bound			0.5006			0.6935			0.7037			0.4455

Source: Authors' calculations from the following sources: Trial Exhibit P623; Tables 4.2, B.3.2, B.4.1, and B.5.1–B.5.6 from Document 415-8; and Tables A.5R and B.5.1R–B.5.6R from Document 415-9. See Appendix B.2.8 for a complete discussion of the calculations.

Table 9: Admissions Probabilities as ALDC or Disadvantaged

	Admission Probability when...				
Applicant Race	Non-ALDC	Legacy	Double Legacy	Dean's List	Disadvantaged
White	0.01	0.079	0.146	0.215	0.049
	0.05	0.309	0.471	0.588	0.213
	0.10	0.486	0.653	0.751	0.364
	0.15	0.600	0.749	0.827	0.476
	0.20	0.680	0.809	0.872	0.563
African American	0.01	0.037	0.072	0.072	0.011
	0.05	0.169	0.288	0.288	0.053
	0.10	0.300	0.460	0.461	0.105
	0.15	0.405	0.575	0.576	0.158
	0.20	0.491	0.657	0.658	0.209
Hispanic	0.01	0.038	0.073	0.067	0.027
	0.05	0.170	0.290	0.272	0.128
	0.10	0.303	0.464	0.440	0.236
	0.15	0.408	0.578	0.556	0.329
	0.20	0.494	0.660	0.639	0.410
Asian American	0.01	0.138	0.242	0.315	0.058
	0.05	0.456	0.625	0.705	0.241
	0.10	0.639	0.779	0.835	0.402
	0.15	0.737	0.848	0.889	0.516
	0.20	0.799	0.888	0.919	0.602

Source: Authors' calculations based on Appendix Table D4 (which in turn is an excerpt of Table B.7.2R of Document 415-9) and using the formula in Equation (1) and discussion in the following paragraph. See Appendix B.2.9 for a complete discussion of the calculations.

Table 10: ALDC Admit Rates when Preferences are Removed

	White	African American	Hispanic	Asian American
Total Admits	4,993	1,392	1,283	2,443
LDC Admits	1,362	81	122	270
Athlete Admits	817	124	54	101
Total Applicants	62,586	16,103	18,383	41,258
LDC Applicants	4,075	295	352	727
Athlete Applicants	929	144	61	116
Non-ALDC Admit Rate	0.0489	0.0758	0.0616	0.0513
ALDC Admit Rate	0.4355	0.4670	0.4262	0.4401
LDC Admit Rate	0.3342	0.2746	0.3466	0.3714
Remove ALDC Preferences				
Admit Rate for Previous Admits	0.6778	0.9164	0.9352	0.8952
Admit Rate for Previous ALDC Admits	0.2617	0.4323	0.5276	0.3099
ALDC Admit Rate	0.1140	0.2019	0.2248	0.1364
LDC Admit Rate, Upper Bound	0.1399	0.3004	0.2638	0.1581

Source: Authors' calculations based on the following sources: Table 3 of Exhibit 287; Equation (2); Table B.3.2 from Document 415-8; and Tables B.3.1R and B.3.2R from Document 415-9. See Appendix B.2.10 for a complete discussion of the calculations.

Table 11: Total Admits by Race under Different Admissions Policies, Expanded Sample

	White	African American	Hispanic	Asian American
Model	4,802	1,367	1,365	2,358
No legacy preferences	4,598	1,423	1,428	2,458
No athlete preferences	4,499	1,366	1,462	2,569
No race/legacy/athlete	4,947	428	792	3,564

Source: Data presented in Panel 2 of Table 8.2R of Document 415-9.

Table 12: Inclusion of ALDC Applicants Distorts Effect of Other Admissions Criteria

	Baseline Coefficients	Expanded Coefficients	Expanded plus Athletes	% Increase over Expanded	% Increase over Expanded plus Athletes
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.

A Description of Legal Documents Used

We list in Appendix Table A1 all publicly released documents pertaining to the *SFFA v. Harvard* trial that we use in our analysis. A complete list of exhibits admitted into evidence during trial is available as Document 611.

Table A1: 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
Day 3 Trial Transcript	Transcript of Day 3 of trial proceedings
Day 5 Trial Transcript	Transcript of Day 5 of trial proceedings
Day 6 Trial Transcript	Transcript of Day 6 of trial proceedings
Trial Exhibit DX 002	2012 Harvard admissions reader casebook
Trial Exhibit DX 005	2013–2014 Harvard alumni interviewer handbook
Trial Exhibit DX 024	2012 Harvard admissions reader casebook discussion guide
Trial Exhibit DX 706	Share of admitted students in ALDC categories
Trial Exhibit DX 730	Academic qualifications and profile ratings of transfer applicants
Trial Exhibit DX 746	Racial composition of admitted ALDC students
Trial Exhibit P001	Class of 2018 application reading procedures
Trial Exhibit P104	Email exchange discussing admission of donors' children
Trial Exhibit P106	Email exchange rating a donor as a "2"
Trial Exhibit P164	Class of 2018 one-pager
Trial Exhibit P316	Report of Harvard's Committee to Study Race-Neutral Alternatives
Trial Exhibit P555	Office for Civil Rights Report (1990)
Trial Exhibit P618	Admit rates for ALDCs and non-ALDCs by academic rating
Trial Exhibit P619	List of applicants interviewed and admitted
Trial Exhibit P621	Ratings frequencies for baseline sample
Trial Exhibit P623	Ratings frequencies for expanded sample
Trial Exhibit P626	Number and share of applicants by academic index
Trial Exhibit P723	Class of 2023 application reading procedures

B Recovering Data from Public Sources

B.1 General Approach

As discussed in Section 2.2, the analysis presented in the current paper is based entirely on publicly available documents since we no longer have access to the underlying application data. This presents a number of challenges since the publicly released reports do not focus on ALDC applicants or admissions. As an example, none of the reports provide summaries of the application characteristics of ALDCs. We are able to overcome these limitations in a number of ways.

For LDC applicants (i.e. legacy, dean's interest list, and children of faculty and staff), we are able to provide detailed and precise analysis of their demographic characteristics, internal Harvard ratings, and academic preparation by exploiting the fact that Document 415-9 reports findings for two samples: one that includes LDC applicants (expanded) and one that excludes these special applicants (baseline). Using the contrast in characteristics and ratings between these two samples we are able to infer information about LDC applicants and admits.[93]

Isolating the characteristics for recruited athletes is more challenging. In Document 415-9, athletes are excluded from both the baseline and expanded sample. However, in Document 415-8, recruited athletes are part of the expanded sample, along with LDC applicants. This suggests that a comparison between the expanded sample across the two reports would identify athlete attributes. However, there are other minor modifications to the expanded sample across the two reports that make direct comparisons impossible.[94] There are 150,701 applicants in the expanded sample in Document 415-8 including athletes, while there are only 148,769 applicants in the expanded sample in Document 415-9 excluding athletes. There are

[93]The calculations for demographic characteristics in Table 4 are based upon shares that have already been rounded. As a result, the numbers we report will have rounding error in them. It is straightforward to show that the rounding bias is sensitive to the ratio of the number of applicants/admits among non-ALDCs to LDCs. Because this ratio will be much smaller for admits, rounding error will be smaller when describing the characteristics of admitted LDCs. Additionally, we do not report average SAT score and academic index among LDC applicants and admits since only two significant digits are included in the original report, exacerbating rounding concerns.

[94]For example, a small group of applicants are dropped from the sample in Document 415-9 as a result of recoding some of the profile ratings as missing.

1,343 recruited athlete applicants, which means that the expanded sample in Document 415-8 has 589 additional applicants that are non-athletes. Simply comparing attributes across the two expanded samples will confound the characteristics of recruited athletes with these additional 589 applicants. If we focus on describing admitted athletes, we can be more precise. Of the 589 non-athlete, extra applicants in Document 415-8, only 64 are admitted.[95] Among the 1,343 recruited athlete applicants, 1,179 are admitted. As a result, if we compare the characteristics of admitted applicants in the expanded sample across the two reports, 95% of the gap is related to recruited athletes.

Our descriptive analysis of athletes follows the strategy outlined above, where we infer recruited athlete attributes by taking the difference across the expanded sample in Document 415-8 and Document 415-9. We account for the presence of non-athletes in this comparison by providing upper and lower bounds for athlete attributes. These bounds are generated by assuming that the non-athlete applicants and admits all have a particular feature or all fail to have a particular feature.

In addition to descriptives for ALDC applicants, we are also interested in understanding the advantages that ALDC applicants receive in the admission process and how the admitted class would change if these preferences were removed. Because we do not have access to the underlying application data, here we must rely on the admissions models and coefficients presented in Document 415-8 and Document 415-9, along with additional calculations contained in Exhibit 287 (Plaintiff declaration).

B.2 Table Details

In this section, we provide the exact steps we take to construct each of the main tables.

B.2.1 Table 1

Table 1 is constructed from numbers contained in Table A.2R in Document 415-9. The sample utilized in constructing this figure consists of all applicants to Harvard, excluding:

[95] None of these 64 admits are recruited athletes as the total number of admitted recruited athletes from Document 415-8 Table B.3.2 matches the number of admitted recruited athletes in Document 415-9 Table A.5R.

(1) transfer and foreign applicants; (2) applications that were withdrawn, incomplete, or departed; and (3) repeat applicants. To see this, note that the total sample size is 166,727 applicants, which matches the remaining observations after these groups are eliminated in Table A.5R of Document 415-9.

B.2.2 Table 2

Table 2 is directly constructed from numbers contained in Trial Exhibit P618.

B.2.3 Table 3

Table 3 is constructed using information from Tables B.3.1R and B.3.2R in Document 415-9 and Table B.3.2 in Document 415-8.

The racial distribution of non-ALDC applicants and admits (Panels B and C) can be determined using Table B.3.1R since we know the total number of applicants and admits by race. The racial distribution of athlete applicants and admits is based on Table B.3.2. We use the share of applicants and admits that are athletes, by race, to calculate the number of athlete applicants and admits by race and then divide by the total number of athlete admits. To avoid rounding issues here, we confirm our numbers for admits with Trial Exhibit DX 746, although we cannot do this for Hispanics because they are combined with other race groups in the exhibit. The racial distribution of LDC applicants and admits is determined using Table B.3.2R. We use the total number of applicants and admits by race along with the share of applicants and admits that are legacies, dean's interest, and child of faculty/staff by race to calculate the number in each group and then divide by the total number in each group. When constructing values for children of faculty or staff, we simply add the two groups together. We know there is essentially no overlap between the groups since the number of admitted faculty children and staff children constructed from Table B.3.2R matches the total number of combined faculty/staff admits in Trial Exhibit DX 746.

Admit rates by race for non-ALDC applicants come directly from Table B.3.1R. We use Table B.3.2 to calculate admit rates for athletes by race. This is based on the total number of admits and applicants along with the share of admits and applicants that are athletes by racial group. We perform a similar calculation for LDC applicants using Table B.3.2R,

where again we simply add the children of faculty and staff together.

Finally, to calculate the proportion of admits that belong to each of the applicant groups by race, we exploit information already accumulated. The total number of admits for each race is given by the totals at the bottom of Table B.3.2R, plus the number of admitted athletes for each race as determined by Table B.3.2. Then, we simply calculate the share of total admits for each race that fall into each applicant group.

B.2.4 Table 4

Table 4 is constructed using Tables B.3.1R and B.3.2R from Document 415-9. The only difference between the two tables is the inclusion of LDC applicants in Table B.3.2R. Thus, to construct Table 5 we calculate the number of disadvantaged students, for example, who are admitted and rejected by race for each table using the available shares and totals. The difference in these values across the two tables then tells us how many LDC admits and rejects are disadvantaged. Using these values plus the total number of LDC admits and rejects, we construct the share of LDC applicants that are disadvantaged. We utilize an identical approach for the other applicant characteristics.

B.2.5 Table 5

Table 5 is constructed using Trial Exhibits P621 and P623 along with Tables B.3.1R and B.3.2R from Document 415-9. Trial exhibits P621 and P623 provide the raw counts of the number of admits, rejects, and total applicants within each rating bin for all of Harvard's internal ratings. P621 includes only non-ALDC applicants, while P623 includes non-ALDC and LDC applicants. Thus, taking the differences between the raw counts across trial exhibits tells us the number of LDC admits, rejects, and applicants within each rating bin. When constructing the share of LDC applicants receiving higher than a 3+, we use as the denominator the total number of LDC admits, rejects, and applicants, which we know from Tables B.3.1R and B.3.2R. The corresponding numbers for non-ALDC applicants can be constructed using P621 and Table B.3.1R.

B.2.6 Table 6

Table 6 is constructed using Tables 5.1R, 5.2R, B.5.1R and B.5.2R from Document 415-9. The share of non-ALDC applicants in each decile is taken directly from Table 5.1R. The share of LDC applicants in each decile is calculated by taking the difference in the raw counts across Tables 5.1R and B.5.1R, since the only difference is the inclusion of LDC applicants in the latter table. We then divide the number in each decile by the total number of LDC applicants by race. The admit rates for non-ALDC applicants in each decile come directly from Table 5.2R. For LDC applicants, we use the admit rates along with the totals in each decile to calculate the number of admits in each decile for the non-ALDC sample and the sample that also includes LDC applicants. Taking the difference provides the number of LDC admits in each decile, which we then divide by the total number of LDC applicants in each decile.

B.2.7 Table 7

Table 7 is constructed using Tables 5.1R, 5.2R, B.5.1R and B.5.2R from Document 415-9. The total number of admits in each decile is calculated by taking the total number of applicants in each decile from Table B.5.1R and multiplying by that admit rate in each decile from Table B.5.2R. The total number of non-ALDC admits can be generated in a similar manner using Tables 5.1R and 5.2R. The number of LDC admits in each decile is simply the difference between the total admits and non-ALDC admits in each decile.

B.2.8 Table 8

Table 8 is constructed using Trial Exhibit P623, Tables 4.2, B.3.2, B.4.1, and B.5.1–B.5.6 from Document 415-8, and Tables A.5R and B.5.1R–B.5.6R from Document 415-9.

The challenge in constructing ratings for athletes is that these numbers are never directly reported in any publicly released document, and unlike for LDC applicants, there are no two samples we can directly compare that will describe only athletes. The closest we come is to compare the ratings distribution for the expanded sample in Document 415-8 with the ratings distribution for expanded sample in Document 415-9 since the former includes

athletes while the latter does not. However, differences in ratings between these two samples does not identify only athletes as there are additional sample changes to account for.

Relative to the rebuttal expanded sample in Document 415-9, the original expanded sample in Document 415-8 contains athletes and additional applicants with missing profile ratings. There are respectively 1,343 and 1,179 athlete applicants and admits in the original expanded sample (verified from Table B.3.2) that are excluded from the rebuttal expanded sample (verified in Table A.5R). In addition, the original expanded sample includes 64 admits and 589 applicants that are not athletes and not included in the rebuttal expanded sample (Add the athlete numbers to the expanded sample totals in Table B.3.2R and examine the difference between these and the expanded sample totals in Table B.3.2). We can then use the change in the number of admits and applicants receiving each Harvard rating across reports to determine the rating distribution for the composite group of athletes and applicants missing ratings.

There is one additional complication to the above procedure. While we know precisely the number of admits and applicants in each rating group for the rebuttal expanded sample (see Trial Exhibit P623), we do not know the corresponding numbers for the original expanded sample. Table B.4.1 provides the shares in each rating group for the original expanded sample, but for some of the ratings we exclude missing ratings when calculating the share (we do the same thing in Table B.4.1R in the rebuttal, but because we have P623 we can overcome this). Thus, we cannot directly calculate the number of individuals in each bin since we do not know the precise denominator. We overcome this with a multi-step process.

First, the academic index decile tables (B.5.1-B.5.6 and B.5.1R and B.5.6R) provide information on the share of applicants receiving a 2 or higher on each rating, including the missing. If this share represented all of the applicants, we could use this number to infer for each rating how many applicants are excluded because they lack a valid rating. There is one small correction that needs to be made first, which is that the decile analysis excludes individuals whose GPAs appear incorrect. However, because P623 tells us the actual number receiving a 2 or higher in the rebuttal expanded sample, we can determine how many of the excluded GPA group have a 2 or higher. Because the number of missing GPA individuals is essentially identical across the original and expanded samples (compare the totals in Tables

B.5.1 and B.3.2 with the totals in Tables B.5.1R and B.3.2R), we also know the distribution of 2 or higher ratings for missing GPA applicants in the original expanded sample. In other words, none of the missing GPA applicants are athletes or missing ratings such that they are eliminated in the rebuttal expanded sample. Using the fact that we know the number of invalid GPA excluded from the decile analysis and the share of these that receive a 2 or higher for each rating, we can calculate the share of applicants receiving a 2 or higher on each rating, including the missing for the original expanded sample.

With this knowledge we can directly calculate the total number of applicants with an invalid rating that are excluded from the calculations in Table B.4.1 and 4.2 focused on all applicants. Note that Table 4.2 includes the overall rating for the expanded sample in the original report. However, we still do not know the number of invalid ratings that are rejects and admits. Appendix B.3 provides the formulas for calculating these values using knowledge of the total number of missing, along with share of admits receiving a 2 or higher excluding the invalid ratings.

Once we know the total number of applicants and admits with invalid ratings for each category, we can determine the total number of individuals receiving a 2 or higher on each rating for the expanded sample in the original report. Note that for the Overall, Academic, Personal, and Extracurricular ratings, we do not adjust for missing since the share receiving a 2 or higher from the decile tables (Tables B.5.3-B.5.6) matches almost exactly the shares reported in Table B.4.1. This indicates that there are no applicants with invalid ratings, which is consistent with the sample selection criteria described in Table A.5.

The above procedure essentially allows us to create an analog to Trial Exhibit P623 for the expanded sample in the original report. We can then easily determine the number of athletes and missing rating applicants and admits with a 2 or higher for each rating by taking the difference in the totals across the two samples. Once we have this number we proceed to generate lower and upper bounds for the share of athlete applicants and admits receiving a 2 or higher by either assuming all of the missing rating applicants and admits received a 2 or higher, or that none of them received a 2 or higher. For non-ALDC applicants and admits, the numbers come from Table 6.

B.2.9 Table 9

Table 9 is constructed using Model (5) from Table B.7.2R of Document 415-9. This table provides coefficient estimates for the impact that being a legacy, double legacy, dean's interest list, and disadvantaged applicant has on the probability of admission. When performing the admission probability transformations, we focus on transformations in the base year. The formula for the transformations is provided in the body of the paper.

B.2.10 Table 10

Table 10 is constructed using the "Full Sample" panel of Table 3 of Exhibit 287, Table B.3.2 of Document 415-8, and Tables B.3.1R and B.3.2R of Document 415-9. The top half of the table is constructed using information from the tables in the reports to construct total admits, LDC admits, athlete admits, and the corresponding application numbers. Exhibit 287 provides information on the admit rate for previous admits when ALDC preferences are removed. Using this number, along with the top of the table, we can construct the remaining numbers.

B.2.11 Table 11

The numbers in Table 11 are drawn directly from Table 8.2R in Document 415-9. They are in the second panel for the preferred model.

B.2.12 Table 12

Table 12 is a replica of Table 2 from Exhibit 287.

B.3 Missing Observations in the Ratings Analysis

We assume that, from the decile analysis, we see $\overline{X}$, the fraction who received a 2 or better for some X, and the corresponding number of observations, N. However, in the ratings analysis, there are some applicants who may be missing a rating, N_m, and this number is unobserved. We know that this rating will be classified as worse than a 2 in the decile

analysis. The ratings analysis gives $\overline{X}^*$. The relationship between $\overline{X}$ and $\overline{X}^*$ is given by:

$$\overline{X} = \overline{X}^* \frac{(N - N_m)}{N} \tag{B.1}$$

Solving for N_m yields:

$$N_m = N - \frac{N\overline{X}}{\overline{X}^*} \tag{B.2}$$

We now want to solve for the share of the N_m that have been admitted versus rejected. Let ω^R give the fraction of N_m that rejected. We observe $\overline{X}^{*R}$ and $\overline{X}^{*A}$, the share receiving a 2 or higher among rejects and admits when those missing the rating are excluded.

We then use the following relationships to recover ω^R, $\overline{X}^A$, and $\overline{X}^R$:

$$\overline{X}^A = \overline{X}^{*A} \frac{(N^A - (1 - \omega^R)N_m)}{N^A} \tag{B.3}$$

$$\overline{X}^R = \overline{X}^{*R} \frac{(N^R - \omega^R N_m)}{N^R} \tag{B.4}$$

$$\overline{X}^* = \frac{(N^A - (1 - \omega^R)N_m)\overline{X}^{*A} + (N^R - \omega^R N_m)\overline{X}^{*R}}{N - N_m} \tag{B.5}$$

One we have these for both the baseline and expanded samples, we can partition out the ratings for ALDCs where missing ratings are present.

Using the third equation, we can solve for ω^R:

$$\omega^R = \frac{\overline{X}^*(N - N_m) - (N^A - N_m)\overline{X}^{*A} - N^R\overline{X}^{*R}}{N_m(\overline{X}^{*A} - \overline{X}^{*R})} \tag{B.6}$$

C Modeling choices

In this section, we discuss our basic approach to modeling Harvard admissions. The admissions data made available to us as part of the SFFA lawsuit cover six admissions cycles and include hundreds of variables describing each applicant. It is not feasible to include every variable in every year, as we would ultimately have as many regressors as admits. In the paragraphs to follow we briefly discuss some of the key modeling decisions we make that allow us to capture admissions decisions in a simple, yet accurate manner.

The first decision we face is whether to pool the data and estimate a joint admissions model with indicators for admissions cycle, or estimate separate admissions models for each cycle. Our preferred approach is to utilize a pooled model. 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 threshold for admission may change across cycles. If the pool of applicants is simply becoming more competitive over time, meaning that the baseline admit probability is declining, allowing for admission cycle indicators will capture this. However, there may be some applicant characteristics that Harvard seeks to balance within each admissions cycle. As an example, Harvard might target an admitted class where 20% of the students are humanities majors. This would mean that the probability of being accepted as an intended humanity major will vary with the number of other intended humanity major applicants in a given cycle. A pooled admissions model can capture this heterogeneity by including interactions between intended major and indicators for admission cycle. In an admission cycle where there are many intended humanities majors, the interaction coefficient will be negative relative to the baseline humanity effect.

While the admissions impact of certain characteristics is likely to change over time, most applicant characteristics are likely to be valued similarly across admissions cycles. For example, there is little reason to expect that Harvard would value an academic or personal rating of 1 differently from year to year. When employing a pooled model, the question is a matter of knowing which applicant characteristics are likely to have time-varying impacts. Fortunately, during the weeks and months that Harvard is making final admissions decisions,

the admissions office publishes statistics about the makeup of the current admitted class, as well as how these numbers compare to previous classes. Admissions officers can use these "one-pagers" to generate similarly constituted admit classes over time, even if the applicant pool is changing. 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.

The second critical decision when modeling Harvard admissions decisions is which applicant characteristics to include in the model. Our approach in this dimension is to first include variables that Harvard readily admits influence admissions decisions, such as the various internal Harvard ratings, race, and disadvantaged status. Choosing among the potentially hundreds of other variables such as test scores, high school GPA, intended major, and high school and neighborhood characteristics is more challenging. The basic rubric we apply is as follows. First, to be included a variable must be reasonably related to the admissions decision and have no other variable already in the model that captures the same dimension of the applicant. Second, the variable itself cannot be contaminated by other preferences we are seeking to measure. For example, if an applicant's overall rating is influenced by the ALDC status of the applicant, then it is inappropriate to include the overall rating when investigating the impact of ALDC status on admissions. Third, the variable should display consistent patterns over time, an indication that it is a reliable measure. Finally, the variable must be available every admissions cycle, otherwise we cannot include it in our pooled model.

Employing the above criteria results in more than 350 variables being identified as relevant controls in a model of admissions. Yet, there are a few variables we omit from our preferred specification that are worth discussing. Parental occupation and intended career are two variables we exclude from our model. First, both variables have analogs that we include in the model. Parental occupation is essentially a proxy for socioeconomic status, a variable we measure with controls such as disadvantaged status, first generation indicator, and application fee waiver indicator. Intended career is closely tied to intended major. Moreover, both parental occupation and intended career exhibit wild swings across admissions cycles, indicating that they are not consistently measured and thus unreliable. Another vari-

able we exclude from our preferred admissions model is whether the applicant received a staff interview. As illustrated in Trial Exhibit P619, ALDC applicants receive staff interviews at a vastly higher rate that non-ALDC applicants. Thus, a measure of whether an applicant simply receives a staff interview is inappropriate to include in a model aimed at estimating preferences for ALDC applicants. The staff interview indicator itself embeds preferences for ALDC applicants. Finally, Harvard's personal rating is a subjective measure of an applicant's leadership skills and courage. This is a highly contentious variable, since there is ample evidence that the personal rating is heavily influenced by preferences for ALDCs and particular racial groups. As a result, we estimate admissions models with and without this variable.

The final modeling choice we make is which variables to interact in the admissions probability. Conditional on the set of variables identified above, there are potentially thousands of interactions one could generate. Our approach to this issue is driven by knowledge of the college admissions process and past research. For example, previous research finds that the admissions tip associated with being a disadvantaged applicant is different for applicants of different races (Arcidiacono, 2005). As a result, we include interactions between race and disadvantaged. Similarly, racial preferences for applicants are likely to vary according to ALDC status, and thus we include these interactions in the model. The list below describes the full set of variables we include in each of our admissions models. 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 decision, athlete, legacy, double legacy, faculty or staff child, Dean Director's list, 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 decision, race times legacy, and race interacted with an indicator for dean/director's list and/or faculty/staff child

- 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 decision, legacy, dean/director's list or faculty/staff, financial aid, permanent resident, intended major, flag for converted GPA=35, and missing SAT2 average

- Model 5: Model 4 plus indicators for each academic, extracurricular, teacher 1, teacher 2, counselor, alumni personal, and alumni overall ratings, interactions with missing alumni overall rating and race/ethnicity, indicators for whether the applicant had each possible combination of a two or better on Harvard's academic, extracurricular, and athletic profile ratings, indicators for whether the applicant had two or three 2's or better on their school support measures, and an indicator for whether the applicant had 2's or better on both of the alumni ratings

- Model 6: Model 5 plus indicators for each personal rating and indicators for whether the applicant had each possible combination of a two or better on Harvard's profile ratings related to the personal rating

D Supporting Tables

This section reports additional tables in support of the main tables discussed in the exposition.

Table D1: Applicant Ratings by Race and LDC Status

	White		African American		Hispanic		Asian American	
	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC
Overall								
2 or better	4.43	21.60	5.29	19.66	3.88	18.47	4.84	25.58
3	71.39	69.25	42.28	48.47	55.05	69.32	74.44	64.24
4 or worse	24.18	9.15	52.44	31.86	41.07	12.22	20.72	10.18
Academic								
2 or better	45.29	54.43	9.19	15.25	16.74	41.19	60.21	63.27
3	44.83	40.42	40.06	51.19	47.81	47.73	31.80	31.64
4 or worse	9.88	5.15	50.75	33.56	35.44	11.08	7.99	5.09
Extracurricular								
2 or better	24.35	36.22	15.54	30.85	16.83	31.53	28.23	37.83
3	71.94	61.89	76.82	66.10	77.37	64.77	69.78	60.66
4 or worse	3.02	1.62	6.48	2.71	4.79	3.13	1.63	1.38
Athletic								
2 or better	12.79	21.89	6.82	15.93	7.51	18.18	4.81	14.86
3	52.98	52.34	49.28	54.58	48.67	53.41	47.74	52.41
4 or worse	31.65	23.93	39.33	27.12	40.13	24.43	45.35	31.22
Personal								
2 or better	21.27	40.88	19.01	40.68	18.68	38.92	17.64	35.49
3	78.30	58.85	80.52	58.64	80.85	60.51	81.88	64.24
4 or worse	0.43	0.27	0.47	0.68	0.48	0.57	0.48	0.28
Teacher 1								
2 or better	30.42	39.26	17.12	28.47	21.59	31.82	30.79	38.10
3	66.23	59.68	72.04	68.14	70.06	66.76	66.51	60.66
4 or worse	0.52	0.37	0.92	0.68	0.78	0.28	0.47	0.41
Teacher 2								
2 or better	27.13	36.47	14.80	23.05	18.84	32.67	27.41	37.69
3	54.77	51.83	55.01	57.97	55.36	57.67	57.40	52.68
4 or worse	0.37	0.27	0.50	1.02	0.57	0.00	0.41	0.00
Counselor								
2 or better	25.28	37.13	13.86	24.75	16.47	37.22	25.12	36.86
3	69.09	60.74	73.73	71.53	73.59	61.36	70.27	61.49
4 or worse	0.57	0.22	1.62	0.68	1.09	0.28	0.58	0.00
Alumni Personal								
2 or better	49.92	67.85	42.98	58.98	41.39	68.47	50.33	68.50
3	23.63	21.18	24.95	25.76	23.84	19.60	24.24	19.81
4 or worse	5.52	3.78	7.23	4.41	6.77	3.69	6.28	5.50
Alumni overall								
2 or better	36.49	55.34	20.84	37.97	23.61	52.27	40.89	55.98
3	28.02	28.29	24.93	32.54	24.66	29.26	26.38	26.27
4 or worse	13.70	8.66	27.28	17.29	22.06	9.66	12.69	10.59

Source: Authors' calculations from Tables B.3.1R and B.3.2R in Document 415-9 and data presented in Trial Exhibit P621 and Trial Exhibit P623. Numbers do not sum to one since we do not show the share with missing/invalid ratings. 2 or better includes 2−, 2, 2+, and all ratings with a leading 1. 3 includes 3−, 3, and 3+. 4 or worse includes 4−, 4, 4+, and all ratings with a leading 5.

Table D2: Admit Ratings by Race and LDC Status

	White		African American		Hispanic		Asian American	
	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC	Non-ALDC	LDC
Overall								
2 or better	59.70	57.27	59.14	62.96	50.14	50.00	62.36	62.22
3	40.19	42.29	40.86	37.04	49.86	50.00	37.64	37.78
4 or worse	0.11	0.44	0.00	0.00	0.00	0.00	0.00	0.00
Academic								
2 or better	88.77	78.34	59.39	43.21	65.40	70.49	94.40	85.56
3	11.19	20.93	40.52	56.79	34.60	29.51	5.60	14.44
4 or worse	0.04	0.73	0.08	0.00	0.00	0.00	0.00	0.00
Extracurricular								
2 or better	73.03	55.80	51.98	55.56	56.64	50.00	78.28	60.37
3	26.15	43.76	47.18	43.21	42.10	48.36	21.53	39.26
4 or worse	0.11	0.37	0.25	1.23	0.09	1.64	0.00	0.00
Athletic								
2 or better	20.97	27.90	14.24	28.40	15.27	23.77	7.19	18.15
3	44.78	48.02	48.53	41.98	42.82	50.00	43.97	50.74
4 or worse	28.78	22.76	32.60	28.40	35.59	22.95	44.16	29.63
Personal								
2 or better	83.76	70.19	74.39	80.25	77.87	63.93	73.26	60.37
3	16.24	29.66	25.61	19.75	22.13	36.07	26.74	39.63
4 or worse	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00
Teacher 1								
2 or better	77.40	62.11	59.56	65.43	63.50	45.90	74.61	60.00
3	22.60	37.89	40.44	34.57	36.50	54.10	25.34	40.00
4 or worse	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Teacher 2								
2 or better	74.73	57.49	55.69	51.85	64.23	54.10	72.78	53.33
3	22.28	36.78	39.93	34.57	31.71	42.62	23.75	39.26
4 or worse	0.00	0.15	0.00	0.00	0.00	0.00	0.05	0.00
Counselor								
2 or better	76.01	61.97	57.88	58.02	58.36	60.66	73.17	63.33
3	23.03	37.08	41.45	40.74	40.11	39.34	26.11	35.93
4 or worse	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Alumni Personal								
2 or better	91.47	82.01	87.78	80.25	91.42	89.34	91.75	85.19
3	5.76	13.51	8.93	17.28	6.41	5.74	6.27	10.74
4 or worse	0.50	1.17	1.10	1.23	0.27	1.64	0.34	2.22
Alumni overall								
2 or better	86.57	73.94	73.46	62.96	80.13	80.33	90.35	79.26
3	10.20	20.19	21.82	30.86	16.08	15.57	7.43	14.81
4 or worse	0.82	2.28	2.19	4.94	1.81	0.82	0.43	3.70

Source: Authors' calculations from the following sources: Trial Exhibit P623; Tables 4.2, B.3.2, B.4.1, and B.5.1–B.5.6 from Document 415-8; and Tables A.5R and B.5.1R–B.5.6R from Document 415-9. See Appendix B.2.8 for a complete discussion of the calculations. 2 or better includes 2−, 2, 2+, and all ratings with a leading 1. 3 includes 3−, 3, and 3+. 4 or worse includes 4−, 4, 4+, and all ratings with a leading 5.

Table D3: Upper and Lower Bounds on the Share of Applicants and Admits Receiving a 2 or Higher on Various Ratings, by Athlete Status and Race

	White		African American		Hispanic		Asian American	
Rating	Non-ALDC	Recruited Athlete	Non-ALDC	Recruited Athlete	Non-ALDC	Recruited Athlete	Non-ALDC	Recruited Athlete
Panel A: Applicants								
Overall								
Lower bound	0.0443	0.0667	0.0529	0.0000	0.0388	0.0000	0.0484	0.0000
Upper bound		0.2713		0.1667		0.3115		0.3621
Academic								
Lower bound	0.4529	0.0990	0.0919	0.0000	0.1674	0.0000	0.6021	0.0000
Upper bound		0.3036		0.0694		0.3115		0.7759
Extracurricular								
Lower bound	0.2435	0.0000	0.1554	0.0000	0.1683	0.0000	0.2823	0.0000
Upper bound		0.1410		0.2361		0.5082		0.2759
Personal								
Lower bound	0.2127	0.3111	0.1901	0.0000	0.1868	0.0000	0.1764	0.0000
Upper bound		0.5156		0.8056		1.0000		0.5259
Teacher 1								
Lower bound	0.3042	0.2034	0.1712	0.0000	0.2159	0.0000	0.3079	0.0000
Upper bound		0.4080		0.4167		0.8361		0.6034
Teacher 2								
Lower bound	0.2713	0.1507	0.1480	0.0000	0.1884	0.0000	0.2741	0.0000
Upper bound		0.3552		0.3125		0.7705		0.5345
Counselor								
Lower bound	0.2528	0.1830	0.1386	0.0000	0.1647	0.0000	0.2512	0.0000
Upper bound		0.3875		0.3819		0.8361		0.6724
Alumni Personal								
Lower bound	0.4992	0.2820	0.4298	0.0000	0.4139	0.0000	0.5033	0.0000
Upper bound		0.4865		0.7986		1.0000		0.8448
Alumni Overall								
Lower bound	0.3649	0.2024	0.2084	0.0000	0.2361	0.0000	0.4089	0.0000
Upper bound		0.4069		0.5000		0.9344		0.6983
Panel B: Admits								
Overall								
Lower bound	0.5970	0.2705	0.5914	0.1290	0.5014	0.1481	0.6236	0.2376
Upper bound		0.3035		0.1935		0.3333		0.3960
Academic								
Lower bound	0.8877	0.2460	0.5939	0.0000	0.6540	0.1111	0.9440	0.4158
Upper bound		0.2791		0.0645		0.2963		0.5743
Extracurricular								
Lower bound	0.7303	0.0832	0.5198	0.0645	0.5664	0.0000	0.7828	0.0000
Upper bound		0.1163		0.1290		0.1667		0.1188
Personal								
Lower bound	0.8376	0.4676	0.7439	0.6290	0.7787	0.5185	0.7326	0.2871
Upper bound		0.5006		0.6935		0.7037		0.4455
Teacher 1								
Lower bound	0.7740	0.3329	0.5956	0.2339	0.6350	0.2963	0.7461	0.2079
Upper bound		0.3660		0.2984		0.4815		0.3663
Teacher 2								
Lower bound	0.7473	0.2962	0.5569	0.1774	0.6423	0.2593	0.7278	0.2475
Upper bound		0.3293		0.2419		0.4444		0.4059
Counselor								
Lower bound	0.7601	0.3256	0.5788	0.2661	0.5836	0.2593	0.7317	0.2673
Upper bound		0.3586		0.3306		0.4444		0.4257
Alumni Personal								
Lower bound	0.9147	0.3696	0.8778	0.4435	0.9142	0.2963	0.9175	0.3762
Upper bound		0.4027		0.5081		0.4815		0.5347
Alumni Overall								
Lower bound	0.8657	0.3390	0.7346	0.3629	0.8013	0.3704	0.9035	0.3168
Upper bound		0.3721		0.4274		0.5556		0.4752

Source: Authors' calculations from the following sources: Trial Exhibit P623; Tables 4.2, B.3.2, B.4.1, and B.5.1–B.5.6 from Document 415-8; and Tables A.5R and B.5.1R–B.5.6R from Document 415-9. See Appendix B.2.8 for a complete discussion of the calculations.

Table D4: Selected Coefficients, Admissions Model of LDCs and non-ALDCs

	(1)	(2)	(3)	(4)	(5)	(6)
Legacy	1.238	1.650	1.697	1.720	2.141	2.329
	(0.046)	(0.051)	(0.059)	(0.123)	(0.155)	(0.164)
Double Legacy	0.511	0.372	0.377	0.337	0.689	0.738
	(0.090)	(0.101)	(0.101)	(0.106)	(0.130)	(0.135)
Faculty or Staff	1.260	1.410	1.692	1.875	2.472	2.630
	(0.139)	(0.159)	(0.310)	(0.319)	(0.359)	(0.353)
Dean's Interest	1.495	1.931	2.379	2.449	3.301	3.246
	(0.053)	(0.059)	(0.356)	(0.366)	(0.417)	(0.417)
African American	0.486	2.290	2.604	2.815	3.596	3.674
	(0.038)	(0.047)	(0.071)	(0.075)	(0.097)	(0.103)
Hispanic	0.393	1.190	1.271	1.338	1.908	1.959
	(0.037)	(0.042)	(0.061)	(0.064)	(0.081)	(0.086)
Asian American	0.047	-0.400	-0.529	-0.321	-0.389	-0.257
	(0.030)	(0.032)	(0.050)	(0.053)	(0.066)	(0.070)
Disadvantaged	1.172	1.243	1.494	1.616	1.640	1.527
	(0.041)	(0.047)	(0.070)	(0.106)	(0.132)	(0.139)
Legacy × African American			-0.725	-0.716	-0.792	-0.872
			(0.214)	(0.223)	(0.281)	(0.297)
Legacy × Hispanic			-0.536	-0.672	-0.779	-0.736
			(0.183)	(0.192)	(0.235)	(0.240)
Legacy × Asian American			0.398	0.331	0.626	0.612
			(0.142)	(0.150)	(0.187)	(0.195)
Other Special × African American			-0.882	-0.788	-1.261	-1.267
			(0.349)	(0.364)	(0.485)	(0.529)
Other Special × Hispanic			-0.729	-0.692	-1.343	-1.328
			(0.230)	(0.243)	(0.287)	(0.295)
Other Special × Asian American			0.377	0.491	0.515	0.471
			(0.160)	(0.175)	(0.208)	(0.219)
Disadvantaged × African American			-1.023	-1.121	-1.582	-1.565
			(0.104)	(0.108)	(0.135)	(0.142)
Disadvantaged × Hispanic			-0.278	-0.356	-0.618	-0.616
			(0.096)	(0.102)	(0.127)	(0.133)
Disadvantaged × Asian American			0.020	0.023	0.159	0.162
			(0.090)	(0.093)	(0.115)	(0.121)
N	148,769	148,741	148,741	141,701	134,365	134,349
Pseudo R Sq.	0.136	0.294	0.297	0.318	0.555	0.599
Demographics	Y	Y	Y	Y	Y	Y
Academics	N	Y	Y	Y	Y	Y
Race and Gender Interactions	N	N	Y	Y	Y	Y
HS and NBHD Variables	N	N	N	Y	Y	Y
Ratings (excluding Personal)	N	N	N	N	Y	Y
Personal Rating	N	N	N	N	N	Y

Source: Data presented in Table B.7.2R of Document 415-9. All models include year indicators and year indicators interactions.

Table D5: Selected Coefficients, Admissions Model of ALDCs and non-ALDCs

	(1)	(2)	(3)	(4)	(5)	(6)
Athlete	4.487	7.153	7.141	7.245	8.532	7.849
	(0.088)	(0.116)	(0.116)	(0.117)	(0.147)	(0.153)
Legacy	1.244	1.662	1.682	1.658	2.058	1.840
	(0.045)	(0.051)	(0.058)	(0.059)	(0.073)	(0.082)
Double Legacy	0.509	0.370	0.381	0.354	0.607	0.629
	(0.090)	(0.100)	(0.100)	(0.101)	(0.121)	(0.133)
Faculty or Staff Child	1.252	1.389	1.409	1.407	1.822	1.704
	(0.138)	(0.157)	(0.157)	(0.158)	(0.187)	(0.203)
Dean's Interest	1.499	1.941	1.913	1.873	2.307	2.322
	(0.053)	(0.059)	(0.059)	(0.059)	(0.072)	(0.077)
African American	0.420	2.163	2.533	2.622	3.333	2.659
	(0.038)	(0.046)	(0.070)	(0.071)	(0.091)	(0.104)
Hispanic	0.329	1.092	1.170	1.180	1.700	1.419
	(0.038)	(0.043)	(0.063)	(0.065)	(0.080)	(0.091)
Asian American	0.005	-0.438	-0.529	-0.457	-0.436	-0.271
	(0.029)	(0.032)	(0.049)	(0.050)	(0.062)	(0.071)
Female	-0.039	0.250	0.239	0.248	0.145	0.127
	(0.022)	(0.025)	(0.057)	(0.057)	(0.068)	(0.076)
Disadvantaged	1.154	1.224	1.482	1.472	1.364	1.083
	(0.040)	(0.046)	(0.069)	(0.069)	(0.085)	(0.093)
First Generation	0.006	0.170	0.156	0.136	0.074	0.023
	(0.050)	(0.056)	(0.056)	(0.057)	(0.069)	(0.077)
Early Decision	1.611	1.449	1.383	1.384	1.333	1.282
	(0.029)	(0.031)	(0.046)	(0.046)	(0.056)	(0.062)
N	150,701	150,633	150,633	150,587	149,425	144,189
Psuedo R^2	0.187	0.331	0.337	0.343	0.568	0.649
Demographics	Y	Y	Y	Y	Y	Y
Academics	N	Y	Y	Y	Y	Y
Race and Gender Interactions	N	N	Y	Y	Y	Y
HS and NBHD Variables	N	N	N	Y	Y	Y
Ratings (excluding Personal & Overall)	N	N	N	N	Y	Y
Personal & Overall Ratings	N	N	N	N	N	Y

Source: Data presented in Table B.7.2 of Document 415-8.

E Varsity Athletics at Harvard

Harvard fields 42 varsity teams, the most among the nation's NCAA Division I colleges and universities.[96] Nearly 1,200 Harvard undergraduates—or 20 percent of the student body—participate in intercollegiate athletics. In contrast, Stanford's student body (which has about 300 more undergraduates) has 300 fewer athletes: "The Department of Athletics offers 36 varsity sports—20 for women, 16 for men (sailing is a co-ed sport). ... Stanford offers over 350 athletic scholarships and approximately 900 students participate in intercollegiate sports."[97] Harvard has more student athletes than even Ohio State, which supports 1,038 student athletes in 30 different sports[98] with an undergraduate enrollment of nearly 46,000.

Appendix Table E1 contains a complete list of sports offered by Harvard and when Harvard began offering them.

[96] See also Document 419-1, p. 41.

[97] https://facts.stanford.edu/campuslife/athletics/, accessed March 29, 2019.

[98] https://www.collegefactual.com/colleges/ohio-state-university-main-campus/student-life/sports/, accessed April 4, 2019.

Table E1: List of varsity sports offered at Harvard

Academic Year of Inception	Team
1852-1853	Men's Heavyweight Crew
1865-1866	Baseball
1873-1874	Football
1876-1877	Men's Outdoor Track
1881-1882	Men's Lacrosse
1893-1894	Men's Fencing
1897-1898	Men's Ice Hockey
1900-1901	Men's Basketball
1902-1903	Men's Cross Country
1905-1906	Men's Soccer
1913-1914	Wrestling
1921-1922	Men's Lightweight Crew
1922-1923	Men's Indoor Track
1923-1924	Men's Golf
1924-1925	Men's Squash
1925-1926	Men's Tennis
1928-1929	Men's Sailing
1930-1931	Men's Swimming and Diving
1933-1934	Men's Skiing
1970-1971	Women's Heavyweight Crew
1973-1974	Women's Tennis
1974-1975	Women's Basketball
1974-1975	Field Hockey
1974-1975	Women's Lightweight Crew
1974-1975	Women's Outdoor Track and Field
1974-1975	Women's Swimming and Diving
1974-1975	Women's Squash
1975-1976	Women's Fencing
1975-1976	Women's Lacrosse
1976-1977	Women's Cross Country
1976-1977	Women's Sailing
1976-1977	Women's Skiing
1977-1978	Women's Indoor Track and Field
1977-1978	Women's Soccer
1978-1979	Women's Ice Hockey
1980-1981	Softball
1980-1981	Men's Volleyball
1980-1981	Men's Water Polo
1981-1982	Women's Volleyball
1983-1984	Women's Water Polo
1993-1994	Women's Golf
2013-2014	Women's Rugby

Source: https://www.gocrimson.com/information/history/Beginning_Years, accessed March 29, 2019.