Legacy and Athlete Preferences at Harvard

Peter Arcidiacono, *Duke University, National Bureau of Economic Research, and Institute of Labor Economics (IZA)*

Josh Kinsler, *University of Georgia*

Tyler Ransom, *University of Oklahoma and IZA*

We use public documents from the *Students for Fair Admissions v. Harvard University* lawsuit to examine admissions preferences for recruited athletes, legacies, those on the dean's interest list, and children of faculty and staff (ALDCs). More than 43% of white admits are ALDC; the share for African American, Asian American, and Hispanics is less than 16%. Our model of admissions shows that roughly three-quarters of white ALDC admits would have been rejected absent their ALDC status. Removing preferences for athletes and legacies would significantly alter the racial distribution of admitted students away from whites.

We are grateful to Bocar Ba, Peter Blair, Hugh Macartney, Tom Nechyba, Jimmy Roberts, and Juan Carlos Suárez Serrato for helpful comments. Peter Arcidiacono served as an expert witness for Students for Fair Admissions (SFFA) in the *SFFA v. Harvard* case. The SFFA is not funding his work on this paper. Josh Kinsler worked as a consultant for the SFFA in the *SFFA v. Harvard* case. The 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 the 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. Contact the corresponding author, Peter Arcidiacono, at psarcidi@econ.duke.edu. Information concerning access to the data used in this paper is available as supplemental material online.

Submitted September 18, 2020; Accepted January 26, 2021; Electronically published October 25, 2021.

Journal of Labor Economics, volume 40, number 1, January 2022.
 https://doi.org/10.1086/713744

I. 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 also revealed how preferences operate for other 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 estimate the admission advantages ALDC applicants receive relative to typical applicants and show how these advantages impact the racial composition of Harvard's admitted class.

In light 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 United States has been expanding for decades, and college admissions preferences for groups that may already be advantaged are generally viewed poorly (Larson 2006). Additionally, there is widespread concern about the "fairness" of college admissions. 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.[2] Finally, underlying the rising concerns about privilege and fairness in college admissions is the growing competitiveness of the higher education market. Over the past 20 years, application levels have risen dramatically at elite colleges and universities in the United States, with essentially no change in the number of seats available.[3]

While detailed admissions data are tightly guarded by universities, rich data on Harvard admissions were made available as a result of the court case. The 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. The availability of Harvard's internal ratings allows us to better describe the differences between ALDC and typical applicants as well

[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 or relatives have donated or show potential to donate to Harvard.

[2] Legacy preferences are relevant at many competitive schools in the United States. Kochkodin (2019) indicates that children of alumni make up 14%, 22%, 18%, and approximately 12% of the 2022 classes enrolled at Princeton, Notre Dame, the University of North Carolina, and Duke University, respectively.

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

as to account for factors that would typically be unavailable when estimating admissions models.

We find that for each special applicant group under the ALDC umbrella, applicants and admits are disproportionately white and come from higher income households. For example, 40% of typical 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 (i.e., excluding athletes) are stronger than typical applicants. However, the average LDC admit is weaker than the average typical 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 typical admits and, in some cases, typical applicants.[4]

To more precisely measure ALDC admission preferences, we estimate a model of Harvard 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 typical applicant with a 10% chance of admission would see a fivefold increase in admissions likelihood if they were a legacy and more than a sevenfold increase if they were on the dean's interest list; we also find that they would be admitted with near certainty if they were a recruited athlete.[5]

Finally, we explore how the admitted class at Harvard would change if ALDC preferences were eliminated. First, we estimate that roughly one-quarter of white ALDC admits would have been admitted had they been treated as white typical 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.[6] However, the increase in diversity resulting

[4] One might suppose that recruited athletes are a small share of students admitted 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 sec. E of the appendix.

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

[6] 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 nonwhite groups.

from the elimination of legacy and athlete preferences pales in comparison to the diversity benefits stemming from racial preferences. We show that eliminating legacy and athlete preferences *and* racial preferences would result in a 69% and 42% decline in African American and Hispanic admits, respectively.

Because of 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 times, respectively. In our preferred model, the corresponding 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 because of the wealth of data provided,[7] and there is evidence that legacy and athlete preferences have been increasing over time.[8] Hurwitz (2011) uses data from 30 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 nonlegacy 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 to help account for differences between ALDC and typical applicants.

II. Harvard Admissions Data

Our analysis of ALDC applicants and admissions is based on anonymized data provided by Harvard on domestic applicants. The applicants come from the set of students who would have graduated from Harvard in 2014–19. This range includes students who would typically have been applying to Harvard in the fall of 2009–14 and graduating from high school in the spring of 2010–15. In all, the sample consists of 166,727 domestic, nontransfer, complete applications.[9]

[7] The pseudo R^2 values 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. See Norton and Dowd (2018) for a discussion of this issue.

[8] Arcidiacono, Kinsler, and Ransom (2019) demonstrate that the admissions advantages athlete and legacy applicants receive at Harvard have increased substantially over the past 20 years. In particular, the admit rate for athlete and legacy domestic applicants relative to the admit rate for nonathlete and nonlegacy domestic applicants has increased from a ratio of 4:1 to 9:1 between the classes of 2000 and 2017.

[9] For further details about the data and sample selection, see secs. 2.2 and 2.3 of document 415-8 and sec. 3 of document 419-141.

For each applicant, the data contain detailed demographic information, geography, intended major, academic performance in a variety of categories, 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 Harvard's athletic rating.

A key advantage of using Harvard's data to study ALDC preferences is the availability of internal ratings like the athletic rating. Admissions staff, also known as "readers," assign each application a set of numerical codes indicating its strength.[10] 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. 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. Each of the ratings is on a 5-point scale, with lower numbers indicating better ratings.

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 section A of the appendix (available online). All documents are publicly available either at the URL in the reference list or on the Public Access to Court Electronic Records website at https://www.pacer.gov/. Fortunately, the publicly available documents provide enough detailed information for us to infer the characteristics of ALDC applicants relative to their typical applicant peers and the preferences afforded to ALDC applicants in the admissions process. Section B of the appendix shows how each number we present is generated based on information in the public record.

III. Characteristics of ALDC Applicants and Admits

ALDC and typical applicants differ significantly in terms of admissions outcomes and traits. We begin by describing admission rates and racial composition across ALDC status. We then examine how Harvard ratings vary with ALDC status.

A. Admission Rates and Racial Composition of ALDCs

Prior research suggests that ALDC applicants receive large admissions preferences. We observe similar patterns at Harvard, with ALDC applicants admitted at a rate above 30% and typical applicants admitted at rate of just 5.5%. Panel A of table 1 breaks down these aggregate admissions rates by

[10] See trial exhibit P001 for additional details regarding Harvard's ratings and admissions process.

Table 1
Admissions Statistics by Race and ALDC Status

	Typical	Athlete	Legacy	Dean's List	Faculty/Staff
	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
	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
	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
	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	.14
Hispanic	86.28	4.21	6.86	4.36	.62
Asian American	84.81	4.13	6.63	5.40	1.60

SOURCE.—Authors' calculations from data presented in tables B.3.1R and B.3.2R of document 415-9. See sec. B.2.1 of the appendix for a complete discussion of the calculations.

NOTE.—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% because there are other racial groups not shown. The rows in panel D sum to more than 100% because some ALDC admits are in more than one category. ALDCs = recruited athletes, legacies, those on the dean's interest list, and children of faculty and staff.

ALDC category and race.[11] All ALDC groups experience higher admission rates relative to typical applicants, and the pattern is consistent across race. For example, white and African American applicants on the dean's list have admit rates of 42% and 33%, respectively, while their typical applicant counterparts have admit rates of 5% and 8%. Admission rates are especially high for athletes at more than 86% for each racial group.

While all racial groups see higher admit rates as ALDCs, they do not benefit equally from ALDC preferences as the share of each racial group that is ALDC varies considerably. This is illustrated in panels B and C of table 1, which show the racial distribution of applicants and admits for typical applicants and each of the ALDC categories. For each ALDC category, the white shares of applicants and admits are substantially higher than the corresponding

[11] Note that the ALDC categories are not mutually exclusive. For example, it is possible for an applicant both to be a legacy and to be in one of the other ALDC categories.

white shares of non-ALDC applicants and admits. For example, more than 68% of recruited athletes, legacies, and dean's interest list applicants are white, yet less than 41% of typical applicants are white. All other racial groups see higher representation among typical applicants and admits than in any of the corresponding ALDC applicant and admit categories.[12]

The final panel of table 1 documents the share of admits who are in each of the ALDC categories by race. For nonwhites, the share of ALDC admits is less than 16% for each racial/ethnic group.[13] The corresponding share for whites is much higher at more than 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.

B. Harvard Ratings

The high admission rates for ALDC applicants may reflect their relative strength on the dimensions Harvard values. We investigate this by measuring applicant strength as the rate of receiving a score of 2 or better on Harvard's internal ratings. Table 2 shows ratings distributions for Harvard's overall and profile ratings by race for typical and LDC applicants as well as for typical, LDC, and athlete admits.[14] We are only able to bound the ratings distributions for admitted athletes given what is available in the public record.[15]

There are three broad patterns illustrated in table 2. First, LDC applicants are stronger on average than typical applicants, although the relative strength of LDC applicants 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 2s at 18.5% compared with LDC applicants of other racial groups. But this share is 3.5 times larger than the highest share for typical applicants (African Americans at 5.3%). In contrast, on the academic rating, typical Asian American applicants have a higher share of 2s than the LDC applicants of the other three races/ethnicities.

Second, the comparative advantage of LDC applicants over typical applicants generally reverses when we look only at those applicants who were

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

[13] Because of overlap in ALDC group membership, we compute this as 100 minus the typical share.

[14] Tables D1–D3 show the distribution of all Harvard ratings for LDC applicants, LDC admits, and recruited athletes, respectively. Note that table D3 includes recruited athlete applicants. In the main text we focus on recruited athlete admits, since it is more difficult to recover Harvard ratings for recruited athlete applicants from the publicly disclosed documents. This is not particularly limiting, since the admit rate for recruited athletes is above 85%.

[15] Recruited athletes were included in some results in document 415-8 but were completely excluded from document 415-9. See sec. B of the appendix for complete details.

Table 2
Shares (%) of Applicants and Admits Receiving a 2 or Higher on Harvard Ratings

	Applicants		Admits		
	Typical	LDC	Typical	LDC	Athlete
	A. Overall Rating				
White	4.43	21.60	59.70	57.27	(27.05, 30.35)
African American	5.29	19.66	59.14	62.96	(12.90, 19.35)
Hispanic	3.88	18.47	50.14	50.00	(14.81, 33.33)
Asian American	4.84	25.58	62.36	62.22	(23.76, 39.60)
	B. Academic Rating				
White	45.29	54.43	88.77	78.34	(24.60, 27.91)
African American	9.19	15.25	59.39	43.21	(.00, 6.45)
Hispanic	16.74	41.19	65.40	70.49	(11.11, 29.63)
Asian American	60.21	63.27	94.40	85.56	(41.58, 57.43)
	C. Extracurricular Rating				
White	24.35	36.22	73.03	55.80	(8.32, 11.63)
African American	15.54	30.85	51.98	55.56	(6.45, 12.90)
Hispanic	16.83	31.53	56.64	50.00	(.00, 16.67)
Asian American	28.23	37.83	78.28	60.37	(.00, 11.88)
	D. Personal Rating				
White	21.27	40.88	83.76	70.19	(46.76, 50.06)
African American	19.01	40.68	74.39	80.25	(62.90, 69.35)
Hispanic	18.68	38.92	77.87	63.93	(51.85, 70.37)
Asian American	17.64	35.49	73.26	60.37	(28.71, 44.55)
	E. Athletic Rating				
White	12.79	21.89	20.97	27.90	100
African American	6.82	15.93	14.24	28.40	100
Hispanic	7.51	18.18	15.27	23.77	100
Asian American	4.81	14.86	7.19	18.15	100

SOURCE.—Authors' calculations for typical and LDC applicants and admits are from data presented in trial exhibit P621 and trial exhibit P623. See sec. B.2.2 of the appendix for a complete discussion of the calculations. Authors' calculations for athletes are from the following: 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 sec. B.2.2 of the appendix for a complete discussion of the athlete calculations.

NOTE.—Numbers for athlete admits can only be bounded. See footnote 15 in the text. LDC = legacies, those on the dean's interest list, and children of faculty and staff.

admitted. The third and fourth columns of table 2 show that LDC admits are weaker than typical admits in most dimensions. This is not true for the athletic rating, which especially favors LDC applicants.[16]

[16] There are a handful of other ratings and race combinations for which LDC admits are stronger than typical admits, but this occurs only 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 typical applicants. We discuss this further in sec. IV.

Third, recruited athlete admits are universally weaker than typical and LDC admits apart from the athletic rating. For some race and rating combinations, the differences are especially striking. At most, 28% of white athlete admits receive a 2 or higher on the academic rating. In contrast, 89% of white typical admits receive a 2 or higher on the academic rating. In many cases—and in contrast to LDC admits—recruited athlete admits are substantially weaker than typical applicants.[17] Typical applicants on average have higher academic and extracurricular ratings than admitted recruited athletes of the same race for all groups.[18] On the overall and personal ratings, however, recruited athlete admits do have higher scores than typical applicants.

The patterns described above suggest that LDC applicants to Harvard are doubly advantaged in Harvard's holistic admissions process. First, Harvard's extracurricular, personal, and athletic ratings strongly favor these groups (see also Lee and Ries 2019). This is not surprising given the economic advantages LDC applicants experience relative to their typical applicant peers. Harvard application readers label only 1.8% of LDC applicants as disadvantaged, while 12.6% of typical applicants are tagged as disadvantaged.[19] Somewhat surprisingly, a similar pattern emerges for recruited athletes. At most, 3.2% of white admitted athletes are economically disadvantaged, while the corresponding number for white typical admits is 14.6%.[20] Additionally, the *Harvard Crimson* surveys each incoming class about their family background. For the class of 2019, 40.7% of legacy respondents and 26% of recruited athlete respondents respectively have parents who earn more than $500,000, which is the top 1% of US income.[21] The corresponding share for all respondents—including

[17] Indeed, the only applicants admitted with a 5 on the academic rating were recruited athletes. For those who scored a 4 on the academic rating, the admit rate was 0.02% for typical applicants, 3.5% for LDC applicants, and more than 79% for recruited athletes. See trial exhibit P618.

[18] The one exception is for Hispanics on the academic rating, where the share of typical applicants who receive a 2 or better falls within the bounds for recruited athlete admits.

[19] Authors' calculations from data presented in tables B.3.1R and B.3.2R of document 415-9.

[20] These numbers can be derived using table B.3.2 from document 415-8 and table B.3.2R from document 415-9. A similar pattern exists for the other racial groups—the upper bound for athletes is well below the value for typical admits. We focus on white athletes, since they account for close to 70% of admitted recruited athletes (see panel C of table 1).

[21] See Klein and Keto (2015). The share of legacies whose parents earn above $500,000 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 $500,000 according to

$$\Pr(\text{income} > \$500{,}000|\text{legacy}) = \frac{\Pr(\text{legacy}|\text{income} > \$500{,}000) \times \Pr(\text{income} > \$500{,}000)}{\Pr(\text{legacy})}.$$

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

legacies and athletes—is only 15.4%.[22] Recruited athletes at Harvard tend to be advantaged and disproportionately white in part because of the varsity sports Harvard offers, including fencing, sailing, and skiing. These sports are expensive to play and are offered only at elite public and private secondary schools.

Nonacademic factors, such as the personal and extracurricular ratings, 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 elite US universities—opens doors for less advantaged applicants. Harvard's nonacademic ratings appear to achieve that goal only if race is the central characteristic defining disadvantage. The gap between white and underrepresented minority applicants is much smaller for nonacademic ratings compared with the academic rating. However, within racial groups these holistic admissions criteria favor advantaged ALDC applicants. For example, ALDC applicants score substantially better on the personal rating than on the academic rating compared with typical applicants.

The second advantage experienced by ALDC applicants is related directly to admissions decisions. Despite their strength in the applicant pool, the average ALDC admit is weaker than the average typical admit. This suggests a significant preference for ALDC applicants in the admissions process. However, these descriptive statistics cannot rule out the possibility that ALDC 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. We find evidence of large and significant admissions advantages.

IV. Estimates of the Legacy and Athlete Advantage

To better understand the source of the gap in admission rates between ALDC and typical 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 section C of the appendix.[23] Here, we provide a brief overview.

A. Admissions Model and Estimates

The Harvard admissions data cover six admissions cycles and include hundreds of variables describing each applicant. It is not feasible to include every

[22] Remarkably, using data from 20 years earlier, 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). Chetty et al. (2017a) show that the share of Harvard students from the top 1% has stayed steady at about 15% since 2000.

[23] Additional details are provided in sec. 3.7 of document 415-8 and sec. 8 of document 415-9.

variable separately in every year, as we would ultimately have as many regressors as admits. We choose to pool the application cycles and estimate a single logistic regression with indicators for each admissions cycle. Estimating the model with indicators for each admissions cycle ensures that the predicted number of admits matches the actual number of admits. We also choose to exclude athletes. With the admit rate of athletes at 86% compared with less than 5.5% for typical applicants, what variables matter and how they matter is likely different for athletes.[24]

Including only the 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 composition of the admitted class. During the weeks and months that Harvard makes final admissions decisions, the admissions office publishes statistics about the makeup of the provisional admitted class and how these numbers compare to previous classes.[25] Admissions officers can use these "one-pagers" to generate similarly constituted admit classes over time, even if the applicant pool is changing.[26] 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, geographical region, intended major, and academic aptitude, among others. We incorporate many of Harvard's internal ratings, including the academic, extracurricular, and athletic ratings; 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 the overall rating or the personal rating. The overall rating is specifically designed to incorporate admissions preferences, and there is empirical evidence that the personal rating is influenced by LDC preferences.[27] While there is little evidence that Harvard's other ratings incorporate significant bonuses for LDC applicants,

[24] In sec. V, we show that factors such as the academic rating and extracurricular rating become less important when athletes are included in the model. Although we believe it is more appropriate to exclude athletes, we present in the appendix estimates of models with athletes included. It is important to note that the control variables in the models with athletes differ slightly from those in the models excluding athletes. Full details on these differences are provided in sec. 8 of document 415-9.

[25] See trial exhibit P164 for the characteristics that Harvard tracks as they make admissions decisions.

[26] 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.

[27] Document 415-9 estimates an ordered logit model of the personal rating that includes other applicant attributes and finds that legacies receive a large and statistically significant bonus (see table B.6.7R). Despite being an improper control, we include the personal rating in some specifications as a robustness check.

our estimates of the size of LDC preferences would likely be understated if they did.[28]

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.[29] In total, our preferred model includes more than 350 variables.

Our estimated LDC preferences can be interpreted as causal as long as LDC status is uncorrelated with unobserved factors affecting admissions and the admissions index is specified appropriately. We believe the first assumption is reasonable given the richness of Harvard's applicant data. Moreover, it is difficult to envision an alternative method that would allow researchers to exogenously manipulate LDC status. For example, an audit-type study is not feasible when the application process is so data and time intensive.[30]

The second assumption is likely violated since we do not interact LDC status with all applicant attributes in our admissions index. The relative importance of other applicant characteristics, such as academics, matter differently for LDC applicants. For example, no white, Hispanic, or Asian American typical applicants were admitted in the bottom decile of academic preparation, as measured by SAT scores and high school grade point average (GPA).[31] Yet white LDC applicants in the bottom decile of academic preparation were admitted at a higher rate (6.35%) than the average across all typical applicants (5.46%).[32] As a result, a better approach would entail estimating an admissions model using only typical applicants and then applying the coefficients from

[28] For Harvard's other internal ratings, any bumps legacies receive tend to be small or statistically insignificant. See tables B.6.5R and B.6.6R in document 415-9. Note that no model of the athletic rating is available, and thus we cannot say whether the observed LDC advantage in this rating reflects preferences.

[29] 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 table B.3.2R of document 415-9.

[30] If colleges and universities decide to eliminate LDC preferences in the future, it may be possible to exploit such natural experiments. For example, Johns Hopkins University announced in early 2020 that it had begun phasing out legacy admissions in 2014 (Jump 2020). Since 2009, the percentage of legacies at Hopkins in the incoming freshman class has dropped from 12.5% to 3.5%. However, this alone is not sufficient to understand the size of legacy preferences. First, the underlying applicant pool is likely changing over time, complicating any pre-post comparisons. Second, a school that eliminates legacy preferences may respond by reweighting other applicant attributes. This will tend to understate the true extent to which schools value legacy applicants.

[31] Author calculations from data presented in tables 5.1R, 5.2R, B.5.1R and B.5.2R of document 415-9.

[32] Ibid.

this model to generate predicted admission probabilities for LDC applicants. Although we are unable to pursue this alternative, it is likely that our estimated LDC preferences are biased downward as a result of misspecification. When we fail to interact academics with LDC status, for example, the estimated impact of academics is attenuated relative to the truth. Accordingly, LDC preferences do not need to be as strong to rationalize their admissions outcomes. We discuss this issue in more detail in section V.B.

Table 3 displays a subset of the coefficients from our admissions model as we vary the set of controls. Model 5 is our preferred model, yielding a pseudo R^2 of 0.56.[33] Since we interact variables such as legacy and dean's interest with year, all coefficients are for the base year, the class of 2014.[34] Note that even if we exclude Harvard's ratings (model 4) or include the personal rating (model 6), LDC preferences remain large and statistically significant.

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. The odds ratio for legacy is 8.5, and it is even larger for double legacies, those on the dean's interest list, and children of faculty. In a slightly altered model that includes athletes, the odds ratio for athletes exceeds 5,000 (see table D4; tables D1–D4, E1 are available online).[35]

We emphasize that the estimated strength of ALDC preferences is not a product of a particular model specification. While we do not believe our preferred model is overfit, we can compare our estimates with models that have fewer controls. A model that removes Harvard's internal ratings has a worse fit (pseudo $R^2 = 0.32$ rather than 0.56) but still estimates substantial preferences: the odds ratio for legacy is more than 5.5, with dean's interest list and children of faculty seeing even larger magnitudes. The corresponding model in table D4 that excludes the internal ratings gives an athlete odds ratio more than 1,400.

[33] By comparison, the models in Espenshade, Chung, and Walling (2004) achieve a pseudo R^2 no larger than 0.20. The pseudo R^2 is a different measure from the R^2, 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. McFadden (1979) indicates that values of 0.2 to 0.4 for the pseudo R^2 represent an excellent fit.

[34] 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 paragraph. The interactions on legacy range from 0.264 (class of 2015) to −0.474 (class of 2019), suggesting that 2014 is fairly representative.

[35] This alternative model contains fewer controls but is otherwise similar to the model outlined in the previous section. See fig. 7.1 in document 415-8 for a detailed description of the controls in the model. The relevant sample is discussed in sec. 2.3.3 of document 415-8. The coefficient on the recruited athlete variable is substantially larger than any of the other previously discussed preferences, regardless of the controls we include.

Table 3
Selected Coefficients, Admissions Model of LDC and Typical Applicants

	(1)	(2)	(3)	(4)	(5)	(6)
Legacy	1.238	1.650	1.697	1.720	2.141	2.329
	(.046)	(.051)	(.059)	(.123)	(.155)	(.164)
Double legacy	.511	.372	.377	.337	.689	.738
	(.090)	(.101)	(.101)	(.106)	(.130)	(.135)
Faculty or staff	1.260	1.410	1.692	1.875	2.472	2.630
	(.139)	(.159)	(.310)	(.319)	(.359)	(.353)
Dean's interest	1.495	1.931	2.379	2.449	3.301	3.246
	(.053)	(.059)	(.356)	(.366)	(.417)	(.417)
African American	.486	2.290	2.604	2.815	3.596	3.674
	(.038)	(.047)	(.071)	(.075)	(.097)	(.103)
Hispanic	.393	1.190	1.271	1.338	1.908	1.959
	(.037)	(.042)	(.061)	(.064)	(.081)	(.086)
Asian American	.047	−.400	−.529	−.321	−.389	−.257
	(.030)	(.032)	(.050)	(.053)	(.066)	(.070)
Disadvantaged	1.172	1.243	1.494	1.616	1.640	1.527
	(.041)	(.047)	(.070)	(.106)	(.132)	(.139)
Legacy × African American			−.725	−.716	−.792	−.872
			(.214)	(.223)	(.281)	(.297)
Legacy × Hispanic			−.536	−.672	−.779	−.736
			(.183)	(.192)	(.235)	(.240)
Legacy × Asian American			.398	.331	.626	.612
			(.142)	(.150)	(.187)	(.195)
Other special × African American			−.882	−.788	−1.261	−1.267
			(.349)	(.364)	(.485)	(.529)
Other special × Hispanic			−.729	−.692	−1.343	−1.328
			(.230)	(.243)	(.287)	(.295)
Other special × Asian American			.377	.491	.515	.471
			(.160)	(.175)	(.208)	(.219)
Disadvantaged × African American			−1.023	−1.121	−1.582	−1.565
			(.104)	(.108)	(.135)	(.142)
Disadvantaged × Hispanic			−.278	−.356	−.618	−.616
			(.096)	(.102)	(.127)	(.133)
Disadvantaged × Asian American			.020	.023	.159	.162
			(.090)	(.093)	(.115)	(.121)
N	148,769	148,741	148,741	141,701	134,365	134,349
Pseudo R^2	.136	.294	.297	.318	.555	.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
High school and neighborhood 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 indicator interactions.

To further put the size of LDC preferences from our preferred model in context, consider a white typical applicant with a baseline probability of admission of 10%. If this applicant were switched to a legacy, holding all other characteristics fixed, the admission probability would rise to 49%.[36] Switching the same typical applicant to a double legacy or dean's list member would increase the likelihood of admission to 65% and 75%, respectively. Yet shifting this typical applicant into the disadvantaged category only increases the admission probability to 36%. Similar calculations of the impact of LDC preferences can be executed for other racial groups. While the broad patterns are the same, the LDC bumps for African American and Hispanic applicants are more muted.

Using the results in table D4, a similar calculation can be made for athletes. A typical applicant with only a 1% chance of admission would see his admission likelihood increase to 98% if he were a recruited athlete. 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.

B. Removing Preferences

The transformation approach discussed in the previous section can tell us only 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 estimated model coefficients), we cannot determine what would happen to the admissions chances of ALDC applicants more broadly if these preferences were removed. However, calculations presented in exhibit 287 allow us to address this question, abstracting from class size concerns. 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 4 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 typical 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

[36] If the baseline probability of admission for a typical applicant is X, we can calculate the index of observables, Z, for this applicant according to the log odds formula, $Z = \ln(X/(1 - X))$. 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 admission probability would then be given by $\exp(Z + 2.141)/(1 + \exp(Z + 2.141))$.

Table 4
ALDC Admit Rates (%) When Preferences Are Removed

	White	African American	Hispanic	Asian American
Admits:				
Total	4,993	1,392	1,283	2,443
LDC	1,362	81	122	270
Athlete	817	124	54	101
Applicants:				
Total	62,586	16,103	18,383	41,258
LDC	4,075	295	352	727
Athlete	929	144	61	116
Admit rate:				
Typical	4.89	7.58	6.16	5.13
ALDC	43.55	46.70	42.62	44.01
LDC	33.42	27.46	34.66	37.14
Remove ALDC preferences:				
Admit rate for previous admits	67.78	91.64	93.52	89.52
Admit rate for previous ALDC admits	26.17	43.23	52.76	30.99
ALDC admit rate	11.40	20.19	22.48	13.64
LDC admit rate, upper bound	13.99	30.04	26.38	15.81

SOURCE.—Authors' calculations based on the following: table 3 of exhibit 287; eq. (1); table B.3.2 of document 415-8; and tables B.3.1R and B.3.2R of document 415-9. See sec. B.2.3 of the appendix for a complete discussion of the calculations.

NOTE.—ALDCs = recruited athletes, legacies, those on the dean's interest list, and children of faculty and staff; LDCs = legacies, those on the dean's interest list, and children of faculty and staff.

makes it possible to show how much of this admit rate gap between ALDCs and the overall population is the result of ALDC preferences.

Exhibit 287 shows the admit rate for previous admits (by race) if ALDC preferences are removed. This calculation follows directly from Bayes's 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

$$\Pr(y' = 1|y = 1) = \frac{\Pr(y = 1|y' = 1)\Pr(y' = 1)}{\Pr(y = 1)}. \quad (1)$$

The first term in the numerator is, by definition, 1: if a preferred applicant was admitted without a preference, the applicant will also be admitted when a preference is in place. The two remaining terms are the model-predicted probabilities without and with the preference.

The bottom part of table 4 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 typical white applicants would be 26%.

Table 4 also shows what this implies for the overall admit rate of ALDC applicants. The admit rate for all white ALDC applicants would fall from 43.6% to 11.4%, a drop of more than 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%.

V. Counterfactual Simulations

While the transformation approach and the exercises utilizing Bayes's rule highlight the importance of ALDC preferences, neither methodology accounts for capacity constraints. If ALDC preferences were 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 constraints.[37] 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.

A. Methodology

Using our estimated model, we calculate an admissions index for each applicant sans legacy preferences. We construct the index by setting the legacy coefficients to zero but keeping all other coefficients the same, including the coefficients on race.[38] This rules out Harvard changing the weights on correlates of legacy status to undo the removal of legacy preferences.[39] We then use

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

[38] For the estimated model to match the racial distribution of the admitted class in every admission cycle, we add 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.

[39] We view this as a reasonable starting point for understanding how the elimination of legacy preferences impacts the composition of the admitted class. Of course, Harvard could alter the weight on other applicant attributes (including race) or even invent a new internal rating to continue to favor legacy applicants. As an example, Antonovics and Backes (2014) find that University of California campuses changed the weight given to SAT scores, high school GPA, and family background in response to California's ban on race-based affirmative action.

these adjusted admissions indices to construct counterfactual admit classes that match the actual number of Harvard admits every year.

One approach would be to rank applicants according to their adjusted admissions index 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 a deterministic function of the index, which is incompatible with the logit model assumptions and with how Harvard evaluates applications.[40] Instead, we further adjust the admissions index of all applicants by a constant such that the average admission probability without legacy preferences matches the average admission probability with legacy preferences. Numerically, we solve for an index adjustment ϕ_t^* in each admissions cycle t, such that

$$\bar{p}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\exp(X_i \hat{\beta}_{\text{NL}} + \phi_t^*)}{1 + \exp(X_i \hat{\beta}_{\text{NL}} + \phi_t^*)}, \tag{2}$$

where $\bar{p}_t$ 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 $\hat{\beta}_{\text{NL}}$ 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 admission 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 admission probabilities.[41]

Even though recruited athletes are not part of our preferred model, we can still use the preferred 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 recruited athletes like any other applicant and calculate a predicted admission probability from the model.[42]

B. Counterfactual Results

The results of this exercise split out by applicant race are displayed in table 5. The first row for each racial group shows the model-generated number

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

[41] We do not model the equilibrium impact on the application margin, since our data include only applicants to Harvard. We also do not model matriculation.

[42] When we eliminate athlete preferences, we change both the athletic rating and the extracurricular rating to 2 for recruited athletes. See p. 9 of app. A in document 415-9 for additional details.

Table 5
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 exhibit 415-9.

of admits aggregated across the 6 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 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.

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 away from legacy and athlete admits.

The last row in table 5 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 are their interactions. The counterfactual shows that the number of admitted African Americans would be a third of what it was when all of 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.

When only legacy or athlete preferences are eliminated, we estimate that the racial composition of Harvard's admitted class changes by a nontrivial amount. Yet we believe these numbers 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 preferences were eliminated (holding fixed racial preferences). 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%.

Table 6
Inclusion of ALDC Applicants and Its Effect on 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	−.952	−.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.
NOTE.—ALDCs = recruited athletes, legacies, those on the dean's interest list, and children of faculty and staff.

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 from how those same characteristics matter for typical applicants. As described earlier, no white, Hispanic, or Asian American typical applicants in the bottom decile of academic preparation were admitted, while 6.35% of LDC applicants in the bottom decile of academic preparation were admitted.

This phenomenon is illustrated in table 6, where we show how the coefficients on the academic and extracurricular ratings change when athletes and then LDC applicants are excluded from the model.[43] 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 typical applicants compared with 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 in a model with LDC applicants included than when estimated only on typical applicants and becomes smaller still if recruited athletes are included. Similar patterns hold for the extracurricular rating, with the coefficients becoming attenuated when LDC applicants and recruited athletes are included.[44]

[43] The source for this table is table 2 of exhibit 287.

[44] 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. When we account for this property of logit models by dividing each coefficient by the academic rating = 1 coefficient, the results in table 6 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.

In our counterfactuals, we avoided the athlete-generated coefficient distortions by excluding them from the admissions model. We were 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 admission 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.

VI. 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. This paper has focused on the substantial preferences ALDC applicants to Harvard 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 typical applicant.

Each of the ALDC preferences primarily benefit white students. More than 43% of white admits are ALDC, compared with 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 more than 16% of white admits. We show that removing legacy and athlete preferences shifts admissions away from white applicants with other racial groups either increasing or staying the same. Also, since ALDC applicants tend to come from privileged backgrounds, it is likely that 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 enhanced amenities for its student body. But given that the beneficiaries of these practices come from quite advantaged backgrounds and the further evidence that these preferences appear to be increasing over time (Arcidiacono, Kinsler, and Ransom 2019), exposing the scope of these practices may lead them to be reevaluated.

References

Antonovics, Kate, and Ben Backes. 2014. The effect of banning affirmative action on college admissions policies and student quality. *Journal of Human Resources* 49, no. 2:295–322.

Arcidiacono, Peter. 2005. Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica* 73, no. 5:1477–524.

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, no. 4:119–46.

Chappell, Bill, and Merrit Kennedy. 2019. U.S. charges dozens of parents, coaches in massive college admissions scandal. NPR, https://www.npr.org/2019/03/12/702539140/u-s-accuses-actresses-others-of-fraud-in-wide-college-admissions-scandal.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2017a. Mobility report card: Harvard University. Working paper, University of California, Berkeley. https://eml.berkeley.edu/~yagan/college/Harvard_University.pdf.

———. 2017b. Mobility report cards: The role of colleges in intergenerational mobility. NBER Working Paper no. 23618, National Bureau of Economic Research, Cambridge, MA.

Day 3 trial transcript. 2018. In *SFFA v. Harvard*. https://git.io/sffaDay3.

DeSilver, Drew. 2019. A majority of U.S. colleges admit most students who apply. Report, Pew Research Center. 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 *SFFA v. Harvard*. 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 *SFFA v. Harvard*. https://docs.justia.com/cases/federal/district-courts/massachusetts/madce/1:2014cv14176/165519/415/2.html. Plaintiff expert witness rebuttal report.

Document 419-141. 2017. In *SFFA v. Harvard*. 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 *SFFA v. Harvard*. 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.

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, no. 5:1422–46.

Exhibit 287. 2018. In *SFFA v. Harvard*. 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, no. 3:480–92.
Jump, Jim. 2020. Ethical college admissions: Legacy preferences. *Inside Higher Ed*, https://www.insidehighered.com/admissions/views/2020/02/18/legacy-preference-debate-features-sound-arguments-both-sides-opinion.
Klein, Mariel A., and Elizabeth C. Keto. 2015. Class of 2019 by the numbers. *Harvard Crimson*, https://features.thecrimson.com/2015/freshman-survey/makeup-narrative/.
Kochkodin, Brandon. 2019. Notre Dame and Baylor admit more legacies than Harvard and Yale. *Bloomberg*, https://www.bloomberg.com/news/articles/2019-03-21/notre-dame-baylor-top-harvard-yale-for-most-legacies-admitted.
Larson, Carlton F. W. 2006. Titles of nobility, hereditary privilege, and the unconstitutionality of legacy preferences in public school admissions. *Washington University Law Review* 84, no. 6:1375–440.
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*, ed. David A. Hensher and Peter R. Stopher, chap. 13. Kent: Croom Helm.
Norton, Edward C., and Bryan E. Dowd. 2018. Log odds and the interpretation of logit models. *Health Services Research* 53, no. 2:859–78.
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. https://www.epi.org/publication/the-new-gilded-age-income-inequality-in-the-u-s-by-state-metropolitan-area-and-county/.
Trial exhibit P001. 2018. In *SFFA v. Harvard*. https://git.io/sffaP001. Harvard reading procedures for class of 2018.
Trial exhibit P164. 2018. In *SFFA v. Harvard*. https://git.io/sffaP164. One-pager.
Trial exhibit P618. 2018. In *SFFA v. Harvard*. https://git.io/sffaP618. Admit rates for ALDCs and non-ALDCs by academic rating.
Trial exhibit P621. 2018. In *SFFA v. Harvard*. https://git.io/sffaP621. Ratings frequencies for baseline sample.
Trial exhibit P623. 2018. In *SFFA v. Harvard*. https://git.io/sffaP623. Ratings frequencies for expanded sample.