A Modeling enrollment instead of admissions

In this appendix, we examine the implications of changing the university's objective function in Section II such that they value characteristics of the enrollees rather than characteristics of the admits. Denote $f_s(x)$ as the probability that an admitted student of ability x and status s would enroll. Our maximization problem then becomes:

$$\max_{c_l,c_n} U(a,b) \text{ s.t. } N_l\left(1-\int_{c_l}\phi_l(x)f_l(x)dx\right)+N_n\left(1-\int_{c_n}\phi_n(x)f_n(x)dx\right)=\overline{N}$$

$$a := N_l\int_{c_l}x\phi_l(x)f_l(x)dx+N_n\int_{c_n}x\phi_n(x)f_l(x)dx \qquad \text{(quality of enrollees)}$$

$$b := N_l[1-\int_{c_l}\phi_l(x)f_l(x)dx] \qquad \text{(total legacy enrollees)}$$

The first order conditions of the Lagrangian with respect to c_l and c_n yield:

$$0 = -\frac{\partial U}{\partial a}N_l c_l\phi_l(c_l)f_l(c_l)-\frac{\partial U}{\partial b}N_l\phi_l(c_l)f_l(c_l)+\lambda N_l\phi_l(c_l)f_l(c_l) \tag{A.1}$$

$$0 = -\frac{\partial U}{\partial a}N_n c_n\phi_n(c_n)f_n(c_n)+\lambda N_n\phi_n(c_n)f_n(c_n) \tag{A.2}$$

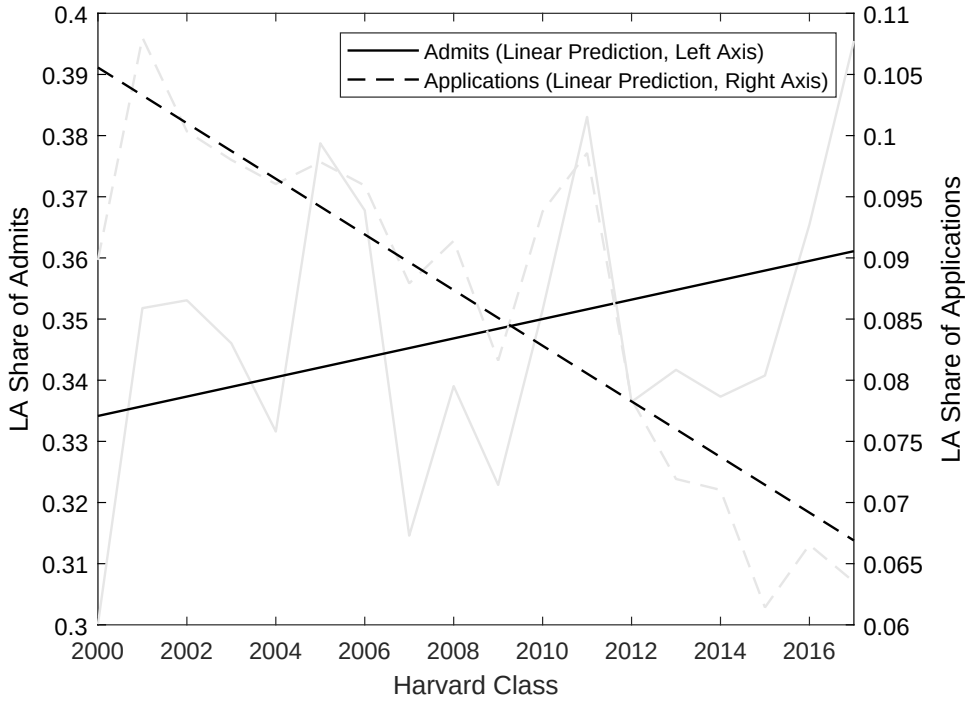
As in the case where the objective function is specified with respect to admissions, combining these two first order conditions yields:

$$c_n-c_l=\frac{\partial U}{\partial b}\left(\frac{\partial U}{\partial a}\right)^{-1} \tag{A.3}$$

implying the discussion in Section II regarding the implications of expanding the number of non-legacy applicants is the same when the objective function includes characteristics of enrollees rather than the characteristics of applicants.

B Supporting Figures and Tables

Figure B1: Trends in LA Share of Applications and Admits for Whites Only



Notes: LA refers to legacies and athletes; NLNA refers to non-legacy, non-athletes. Domestic applicants only.
Source: Authors' calculations from *SFFA v. Harvard* Trial Exhibit DX 042.

Table B1: Domestic Applicants, Admits, and Matriculants by Class

	Legacy and Athlete			Non-Legacy and Non-Athletes		
Year	Applicants	Admits	Matriculants	Applicants	Admits	Matriculants
2000	989	409	324	14,841	1,547	1,203
2001	1,039	485	397	13,242	1,533	1,130
2002	1,026	492	411	13,559	1,478	1,143
2003	1,074	485	409	14,682	1,455	1,117
2004	1,094	491	419	15,108	1,462	1,116
2005	1,101	540	444	15,267	1,414	1,067
2006	1,155	515	437	15,887	1,405	1,060
2007	1,086	444	392	17,049	1,490	1,105
2008	1,038	464	402	15,864	1,469	1,086
2009	1,034	433	381	18,377	1,486	1,104
2010	1,081	473	418	18,240	1,468	1,114
2011	1,121	476	419	17,974	1,439	1,073
2012	1,079	451	405	21,877	1,531	1,080
2013	1,094	472	427	23,556	1,505	1,069
2014	1,095	454	400	24,388	1,557	1,104
2015	1,069	450	399	27,867	1,518	1,068
2016	1,114	449	409	26,861	1,411	1,070
2017	1,094	488	437	27,512	1,349	1,037

Note: Domestic applicants only.
Source: Authors' calculations from *SFFA v. Harvard* Trial Exhibit DX 042.

Table B2: Applicants and Admits by Race and Special Status

	Legacy and Athlete					Non-Legacy and Non-Athlete				
Classes	White	African American	Hispanic	Asian American	Unknown	White	African American	Hispanic	Asian American	Unknown
Panel A: Applications										
2000–2002	1,952	150	66	139	683	17,764	2,728	3,323	10,179	5,901
2003–2005	2,184	135	90	158	651	20,260	3,262	3,780	10,752	5,399
2006–2008	2,367	165	93	171	436	23,445	3,601	4,674	12,600	3,123
2009–2011	2,176	193	116	188	522	21,736	5,559	6,054	14,284	5,292
2012–2014	2,150	187	130	259	500	27,108	8,506	8,556	17,775	6,068
2015–2017	2,243	213	170	320	301	32,940	10,193	10,421	21,591	6,084
Panel B: Admits										
2000–2002	886	59	23	75	328	1,756	515	507	1,007	687
2003–2005	1,033	62	36	88	273	1,901	527	509	880	433
2006–2008	1,055	85	38	64	164	2,043	518	498	1,018	208
2009–2011	946	88	55	73	203	1,748	574	533	1,084	364
2012–2014	912	82	53	104	215	1,783	605	566	1,128	407
2015–2017	946	91	74	151	116	1,641	604	575	1,094	291

Note: Domestic applicants only.
Source: Authors' calculations from *SFFA v. Harvard* Trial Exhibit DX 042.

Table B3: Applicants and Admits by Race and Status after removing non-competitive URM Applicants

	Legacy and Athlete					Non-Legacy and Non-Athlete				
Classes	White	African American	Hispanic	Asian American	Unknown	White	African American	Hispanic	Asian American	Unknown
Panel A: Applications										
2000–2002	1,952	150	66	139	683	17,764	2,728	3,323	10,179	5,901
2003–2005	2,184	135	90	158	651	20,260	3,262	3,780	10,752	5,399
2006–2008	2,367	165	93	171	436	23,445	3,601	4,674	12,600	3,123
2009–2011	2,176	193	116	188	522	21,736	4,365	7,032	14,284	5,292
2012–2014	2,150	187	130	259	500	27,108	5,208	8,391	17,775	6,068
2015–2017	2,243	213	170	320	301	32,940	6,051	9,750	21,591	6,084
Panel B: Admits										
2000–2002	886	59	23	75	328	1,756	515	507	1,007	687
2003–2005	1,033	62	36	88	273	1,901	527	509	880	433
2006–2008	1,055	85	38	64	164	2,043	518	498	1,018	208
2009–2011	946	88	55	73	203	1,748	574	533	1,084	364
2012–2014	912	82	53	104	215	1,783	605	566	1,128	407
2015–2017	946	91	74	151	116	1,641	604	575	1,094	291

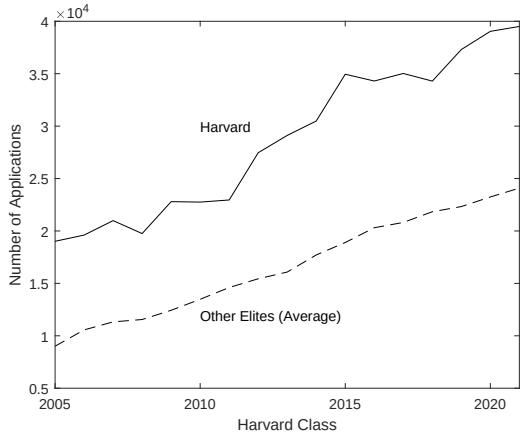
Note: Domestic applicants only. Harvard changed its recruitment of NLNA African American and Hispanic applicants at around the Class of 2008 (see Arcidiacono, Kinsler, and Ransom, 2021). We correct for changes in the applicant pool induced by these new recruiting efforts by assuming that the growth rate in applications for these groups in the post-2008 period has the same ratio as the growth rate for these two groups and Asian American applicants in the pre-2008 period. That is, we compute the number of African American NLNA applicants for t after 2008 as follows:

$$N_{b,t} = \gamma_{0,b} + t \cdot g_a \frac{g_{b,\text{pre-2008}}}{g_{a,\text{pre-2008}}}$$

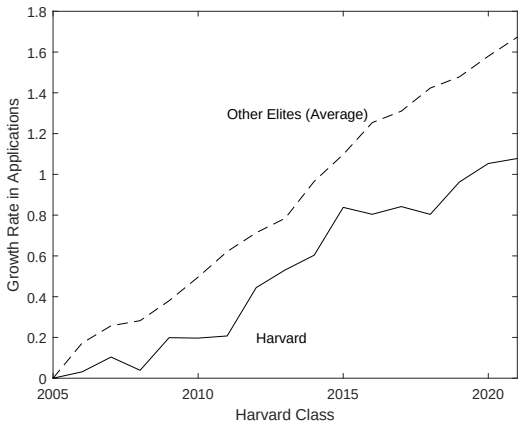
where $\gamma_{0,b}$ is the intercept on a linear trend of all African American NLNA applications (including those induced by additional recruitment efforts), g_a is the linear trend for Asian American NLNA applications over the entire period, and $g_{r,\text{pre-2008}}$ for $r \in \{b,a\}$ are respectively the pre-2008 linear trends in applications for African American NLNA and Asian American NLNA applications. We repeat this process for Hispanic NLNA applicants.
Source: Authors' calculations from *SFFA v. Harvard* Trial Exhibit DX 042.

Figure B2: Application Trends at Harvard and Other Elites

(a) Total Applications



(b) Application Growth



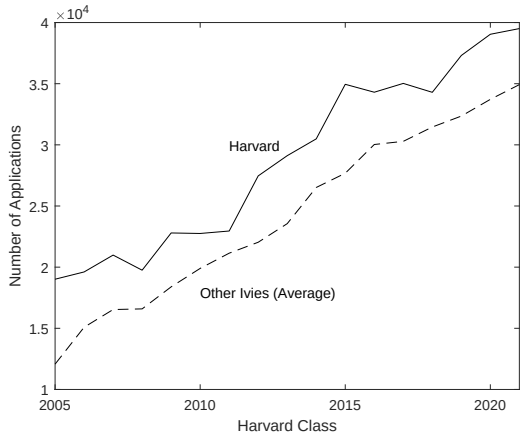
Notes: Panel (a) lists, by year, the total number of applications submitted to Harvard, compared to the total number of applications submitted to Other Elites divided by the number of Other Elite universities. Panel (b) lists growth rates based on the numbers presented in Panel (a).

Other Elites include the following: Amherst College, Caltech, Carnegie Mellon, Columbia, Cornell, Dartmouth, Duke, Harvey Mudd, Johns Hopkins, MIT, Northwestern, Pomona College, Princeton, Rice, Stanford, Swarthmore, Penn, Williams, and Yale. These were chosen because they are 4-year public and private universities that have a 75th percentile math SAT score greater than or equal to 750 between the years of 2001 and 2017, and because they are not missing more than one year of SAT scores or application totals.

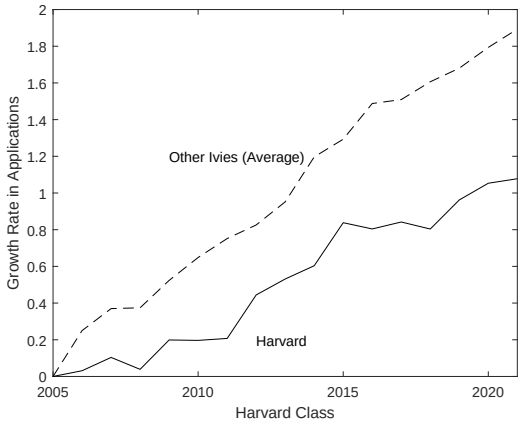
Source: Authors' calculations from *SFFA v. Harvard* Trial Exhibit DX 042 and US National Center for Education Statistics' Integrated Postsecondary Education Data System (IPEDS).

Figure B3: Application Trends at Harvard and Other Ivies

(a) Total Applications



(b) Application Growth



Notes: Panel (a) lists, by year, the total number of applications submitted to Harvard, compared to the total number of applications submitted to other Ivy League institutions divided by the number of other Ivies. Panel (b) lists growth rates based on the numbers presented in Panel (a).

Other Ivies include Columbia, Cornell, Dartmouth, Princeton, Penn, and Yale. Brown is excluded due to incomplete data.

Source: Authors' calculations from *SFFA v. Harvard* Trial Exhibit DX 042 and US National Center for Education Statistics' Integrated Postsecondary Education Data System (IPEDS).

C Counterfactual Academic Ratings Distributions

This appendix describes how we construct counterfactual academic ratings distributions to satisfy a particular ALDC/non-ALDC admit rate ratio. We hold constant the admit rate conditional on receiving a particular academic rating score.

The distribution of ALDC applicants across levels of the academic rating in the current period is available in Table 2. Suppose that in this period the ratio between the average admission probability of ALDC applicants and that of non-ALDC applicants was Y. Suppose that in a prior period it was X, where $X < Y$. Further suppose that the reason for the increase from X to Y was that the ALDC applicants were getting relatively stronger. That is, suppose that the admissions preference for ALDC applicants was fixed.

In this scenario, we show what the academic rating distribution must have looked like for ALDC applicants if all the increases in admission advantage were due to the ALDC applicant pool getting stronger. This requires shifting ALDC applicants across the academic rating distribution while holding fixed the ALDC admission advantage for a given academic rating. We hold constant the distribution of non-ALDC applicants since X and Y are relative comparisons.

Denote $t = 1$ and $t = 0$ as the post and pre periods, respectively. Denote the probability of being admitted in period t given special status s (for ALDCs, $s = 1$) in rating level d by p_{sdt}. Denote the share the share of applicants with a particular rating level as π_{sdt}. Then for the ratio of average admit probabilities to be Y in period 1, the following must hold:

$$Y = \frac{\sum_d \pi_{1d1} p_{1d1}}{\sum_d \pi_{0d1} p_{0d1}} \tag{C.1}$$

Now we want to change the π_{1d1}'s such that the ratio changes to X using a one-parameter

adjustment. Given the π_{1d1}'s, we can express them instead as coming from an ordered logit:

$$\begin{aligned}
\pi_{111} &= \frac{1}{1+\exp(-\kappa_1)} \\
\pi_{1j1} &= \frac{1}{1+\exp(-\kappa_2)} - \frac{1}{1+\exp(-\kappa_1)} \qquad \text{for } 1 < j < D \\
\pi_{1D1} &= 1 - \frac{1}{1+\exp(-\kappa_{D-1})}
\end{aligned} \tag{C.2}$$

where D refers to the top rating level and $\kappa_i < \kappa_j$ for $i < j$.

A shift in the academic distribution involves moving the cutpoints (the κ's) by the same amount. We then look for a Δ such that:

$$\begin{aligned}
\pi^*_{111} &= \frac{1}{1+\exp(\Delta-\kappa_1)} \\
\pi^*_{1j1} &= \frac{1}{1+\exp(\Delta-\kappa_2)} - \frac{1}{1+\exp(\Delta-\kappa_1)} \qquad \text{for } 1 < j < D \\
\pi^*_{1D1} &= 1 - \frac{1}{1+\exp(\Delta-\kappa_{D-1})}
\end{aligned} \tag{C.3}$$

and where the ratio of average admit probabilities is now X:

$$X = \frac{\sum_d \pi^*_{1d1} p_{1d1}}{\sum_d \pi_{0d1} p_{0d1}} \tag{C.4}$$

which after the relevant substitutions is just one equation and one unknown. We can then see how much worse the distribution for ALDC applicants would need to have been relative to non-ALDC applicants to produce a ratio of X rather than Y.

Table C1 illustrates the results of this exercise when $X = 4$, the approximate LA/NLNA admit rate ratio in 2000, and $Y = 8$, the ALDC/non-ALDC admit rate ratio for the classes of 2014-2019. Recall that we assume that the ALDC and LDC counterfactual admit rates by academic rating are identical to the baseline admit rates. This ensures that we hold preferences for these groups fixed as we shift applicants across the academic rating distribution.

The first panel shows that 90% of the ALDC applicants would need to receive an academic rating of 4 or worse to obtain an admit rate ratio of 4:1. This massive shift is driven in part by the high athlete admit rates at the bottom of the academic rating distribution. In the second panel, we assume all athlete applicants face an admission probability of 50%, and find the distribution

of LDC applicants that will generate an overall ALDC/non-ALDC admit rate ratio of 4:1. In this case we find that approximately 40% of LDC applicants would have to receive a 4 or worse on the academic rating.

Table C1: Counterfactual Ratings Distributions

(a) No Restrictions on Athlete Admit Rate

	Typical		ALDC Baseline		ALDC Counterfactual	
Academic Rating	Admit Rate	Share	Admit Rate	Share	Admit Rate	Share
1	66.18	0.43	96.72	0.83	96.72	0.01
2	10.02	41.85	53.17	46.33	53.17	0.92
3	2.40	40.55	35.47	44.22	35.47	9.05
4	0.02	12.73	30.87	7.90	30.87	48.70
5	0.00	4.44	7.41	0.73	7.41	41.32
Weighted Average	5.45		43.61		21.80	
ALDC/Typical Ratio			8.00		4.00	

(b) Set Counterfactual Athlete Admit Rate to 50%

	Typical		LDC Baseline		LDC Counterfactual	
Academic Rating	Admit Rate	Share	Admit Rate	Share	Admit Rate	Share
1	66.18	0.43	96.67	0.99	96.67	0.11
2	10.02	41.85	49.01	51.61	49.01	11.04
3	2.40	40.55	18.09	40.46	18.09	49.12
4	0.02	12.73	3.49	6.17	3.49	33.38
5	0.00	4.44	0.00	0.76	0.00	6.36
Weighted Average	5.45		33.79		15.56	
LDC/Typical Ratio			6.20		2.86	
ALDC/Typical Ratio			8.00		4.00	

Note: Domestic applicants only. Admit rates and shares are expressed as percentages.
Source: Derived from data presented in Trial Exhibit P618 and methodology described in Online Appendix C.