Online Appendix: For Online Publication

The Online Appendix is available at:

https://yungyutsai.github.io/papers/EndowmentTax_OnlineAppendix.pdf

Appendix A: Additional Results

Table A1: List of Colleges Affected by the Net Investment Income Tax

	Student Enrollment		Endowment Assets		Tax Status				
	Total	FTE	Total ($ Million)	Per-student ($ Thousand)	2018	2019	2020	2021	2022
Panel A: Student above 500, and per student Asset above 600K									
Princeton University	8,181	8,082	23,353	2,890	Y	Y	Y	Y	Y
Yale University	12,458	12,383	27,217	2,198	Y	Y	Y	Y	Y
Harvard University	29,908	23,697	37,096	1,565	Y	Y	Y	Y	Y
Stanford University	17,184	16,448	24,785	1,507	Y	Y	Y	Y	Y
Middlebury Institute of International Studies at Monterey	786	717	1,074	1,497	Y	Y	Y	Y	N
Pomona College	1,563	1,558	2,165	1,389	Y	Y	Y	Y	Y
Massachusetts Institute of Technology	11,376	11,247	14,832	1,319	Y	Y	Y	Y	Y
Swarthmore College	1,543	1,542	1,956	1,268	Y	Y	Y	Y	Y
Amherst College	1,849	1,849	2,248	1,216	Y	Y	Y	Y	Y
The Juilliard School	939	872	1,046	1,200	Y	Y	Y	Y	Y
California Institute of Technology	2,240	2,239	2,641	1,179	Y	Y	Y	Y	Y
Williams College	2,150	2,127	2,383	1,121	Y	Y	Y	Y	Y
Grinnell College	1,699	1,672	1,871	1,119	Y	Y	Y	Y	Y
Rice University	6,855	6,662	5,836	876	Y	Y	Y	Y	Y
Cooper Union for the Advancement of Science and Art	964	929	799	860	Y	Y	Y	Y	Y
Bowdoin College	1,806	1,803	1,456	808	Y	Y	Y	Y	Y
Wellesley College	2,482	2,392	1,931	807	Y	Y	Y	Y	Y
University of Notre Dame	12,393	12,256	9,685	790	Y	Y	Y	Y	Y
Dartmouth College	6,409	6,335	4,956	782	Y	Y	Y	Y	Y
Medical College of Wisconsin	1,297	1,178	876	744	Y	Y	Y	Y	Y
Baylor College of Medicine	1,569	1,565	1,134	724	Y	Y	Y	Y	Y
Washington and Lee University	2,160	2,156	1,547	718	Y	Y	Y	Y	Y
University of Richmond	4,131	3,745	2,374	634	Y	Y	Y	Y	Y
Smith College	2,896	2,838	1,767	623	Y	Y	Y	Y	Y
Panel B: Student above 500, and per student Asset between 500 to 600K									
Emory University	14,067	13,009	7,613	585	Y	Y	Y	Y	Y
Claremont McKenna College	1,347	1,346	784	583	Y	Y	Y	Y	Y
Icahn School of Medicine at Mount Sinai	1,203	1,203	675	561	Y	Y	Y	Y	Y
University of Pennsylvania	24,960	22,559	12,213	541	Y	Y	Y	Y	Y
Washington University in St Louis	15,047	13,655	7,215	528	Y	Y	Y	Y	Y
Duke University	15,735	15,218	7,911	520	Y	Y	Y	Y	Y
Bryn Mawr College	1,708	1,661	853	513	Y	Y	Y	Y	Y
Hamilton College	1,883	1,873	955	510	Y	Y	Y	Y	Y
Trinity University	2,466	2,401	1,201	500	Y	N	Y	Y	Y
Panel C: Student above 500, and per student Asset between 400 to 500K									
University of Chicago	15,775	14,136	6,617	468	N	N	N	Y	N
Berry College	2,174	2,115	969	458	N	Y	N	Y	Y
Middlebury College	2,549	2,520	1,074	426	N	N	N	Y	Y
Northwestern University	21,823	18,924	7,948	420	N	N	N	Y	Y
Vassar College	2,424	2,411	1,003	416	N	N	N	Y	N
Colby College	1,879	1,879	775	413	N	N	N	Y	N
Davidson College	1,796	1,796	727	405	N	N	N	Y	Y
Wabash College	842	842	340	404	N	N	N	N	N
Panel D: Student between 400 to 600, and per student Asset above 500K									
Soka University of America	430	430	1,239	2,882	N	N	N	N	N
Principia College	479	479	377	788	N	N	N	N	N

Note: The student enrollment and endowment assets information were in 2016. Full-time equivalent (FTE) is calculated as the sum of full-time and one-third of part-time students. Endowment asset amounts are reported in nominal values. Tax status indicates whether a college is subject to the net investment income tax (NIIT) in a specific year. **Y** refers to being subject to the net investment income tax, while **N** refers to not being subject. The NIIT applies to colleges with over 500 students and more than $500,000 in endowment assets per student.

Table A2: Estimated Net Investment Income Tax Payment

	Average Expenditure / Revenue / Payment ($ Million)				Share of Invest Rev. to Total Rev.	Share of Est. Tax to Total Exp.	Share of Est. Tax to Total Rev.
	Total Expenditure	Total Revenue	Investment Revenue	Estimated NIIT			
Panel A: Student above 500, and per student Asset above 600K							
Princeton University	1,541	3,803	3,073	43.03	58.23%	2.79%	0.82%
Yale University	3,458	6,129	3,400	47.61	43.44%	1.36%	0.61%
Harvard University	4,416	7,412	4,192	58.68	42.82%	1.36%	0.60%
Stanford University	5,176	7,707	3,336	46.70	35.71%	0.91%	0.50%
Pomona College	149	290	216	3.02	47.66%	2.19%	0.67%
Massachusetts Institute of Technology	3,253	5,379	2,997	41.96	40.46%	1.29%	0.57%
Swarthmore College	154	306	235	3.29	52.61%	2.18%	0.74%
Amherst College	194	484	344	4.82	51.52%	2.50%	0.72%
The Juilliard School	98	152	87	1.22	36.80%	1.26%	0.52%
California Institute of Technology	2,822	2,951	304	4.26	9.07%	0.15%	0.13%
Williams College	227	513	355	4.97	50.67%	2.20%	0.71%
Grinnell College	114	327	234	3.27	58.51%	2.96%	0.82%
Rice University	658	1,031	583	8.16	37.45%	1.22%	0.52%
Cooper Union for the Advancement of Science and Art	69	98	69	0.96	67.83%	1.40%	0.95%
Bowdoin College	153	353	256	3.59	50.13%	2.39%	0.70%
Wellesley College	200	404	264	3.70	46.75%	1.92%	0.65%
University of Notre Dame	1,111	2,528	1,674	23.43	43.18%	2.20%	0.60%
Dartmouth College	781	1,460	754	10.55	37.21%	1.38%	0.52%
Medical College of Wisconsin	1,034	1,103	113	1.58	8.20%	0.15%	0.11%
Baylor College of Medicine	1,811	1,838	118	1.65	5.64%	0.09%	0.08%
Washington and Lee University	148	227	130	1.82	36.28%	1.24%	0.51%
University of Richmond	258	401	241	3.37	34.45%	1.32%	0.48%
Smith College	201	340	186	2.60	36.72%	1.39%	0.51%
Panel B: Student above 500, and per student Asset between 500 to 600K							
Emory University	5,581	6,280	853	11.94	12.10%	0.21%	0.17%
Claremont McKenna College	111	229	94	1.32	30.17%	1.27%	0.42%
Icahn School of Medicine at Mount Sinai	2,833	2,980	83	1.17	2.73%	0.04%	0.04%
University of Pennsylvania	9,370	11,344	1,566	21.92	11.95%	0.23%	0.17%
Washington University in St Louis	3,011	4,158	1,435	20.09	23.92%	0.66%	0.33%
Duke University	5,825	7,147	1,707	23.90	17.82%	0.41%	0.25%
Bryn Mawr College	111	186	90	1.26	35.73%	1.18%	0.50%
Hamilton College	124	189	101	1.41	34.89%	1.15%	0.49%
Trinity University	123	203	115	1.61	43.62%	1.31%	0.61%
Panel C: Student above 500, and per student Asset between 400 to 500K							
University of Chicago	3,464	3,869	654	9.15	13.44%	0.26%	0.19%
Berry College	82	138	86	1.20	45.98%	1.47%	0.64%
Middlebury College	237	302	112	1.57	27.86%	0.69%	0.39%
Northwestern University	2,132	2,758	1,055	14.77	28.72%	0.71%	0.40%
Vassar College	171	208	86	1.20	27.74%	0.70%	0.39%
Colby College	141	253	103	1.44	28.47%	1.02%	0.40%
Davidson College	118	223	111	1.55	36.19%	1.29%	0.51%
Wabash College	48	62	22	0.31	23.56%	0.67%	0.33%
Panel D: Student between 400 to 600, and per student Asset above 500K							
Soka University of America	51	124	66	0.92	22.22%	1.89%	0.31%
Principia College	39	62	48	0.67	62.34%	1.77%	0.87%

Note: The data are averaged from 2017 to 2021. Estimated NIIT is calculated by multiplying investment revenue by 1.4%. For observations with negative investment returns, the tax amount is defined as 0. All monetary amounts are adjusted by CPI and reported in 2010 real dollars.

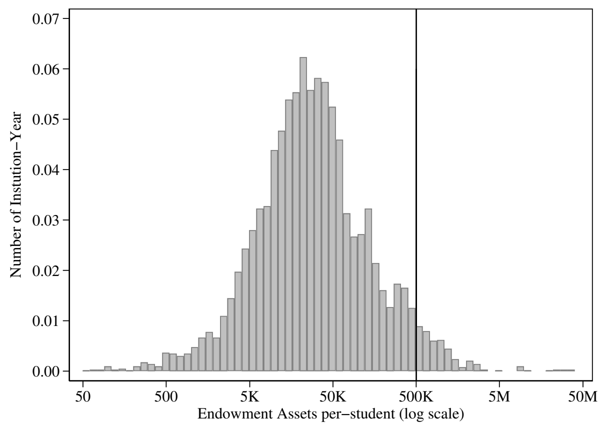
Table A3: Distance of Endowment Assets and Student Enrollment from Tax Threshold

	Distance of from Endowment Threshold				Average Growth Rate	
	Endowment Assets		FTE Enrollment		Endowment Assets	FTE Enrollment
	$ Million	%	Count	%		
Panel A: Student above 500, and per student Asset above 600K						
Princeton University	–19,312	–82.70%	38,625	477.93%	5.36%	0.76%
Yale University	–21,025	–77.25%	42,051	339.59%	6.14%	1.11%
Harvard University	–25,248	–68.06%	50,496	213.09%	2.65%	0.78%
Stanford University	–16,561	–66.82%	33,122	201.37%	7.13%	–0.22%
Middlebury Institute of International Studies at Monterey	–715	–66.60%	1,431	199.44%	1.77%	0.35%
Pomona College	–1,386	–64.01%	2,772	177.89%	4.35%	0.10%
Massachusetts Institute of Technology	–9,209	–62.09%	18,418	163.75%	7.45%	1.28%
Swarthmore College	–1,184	–60.56%	2,369	153.58%	4.69%	0.31%
Amherst College	–1,324	–58.88%	2,647	143.17%	5.71%	0.52%
The Juilliard School	–610	–58.34%	1,220	140.02%	4.59%	–0.11%
California Institute of Technology	–1,521	–57.61%	3,043	135.88%	8.74%	0.50%
Williams College	–1,320	–55.39%	2,640	124.15%	5.43%	0.38%
Grinnell College	–1,035	–55.33%	2,070	123.85%	4.28%	0.49%
Rice University	–2,505	–42.92%	5,009	75.20%	4.63%	2.52%
Cooper Union for the Advancement of Science and Art	–334	–41.84%	669	71.93%	4.64%	–0.92%
Bowdoin College	–555	–38.09%	1,109	61.53%	8.56%	0.44%
Wellesley College	–735	–38.06%	1,470	61.43%	4.28%	–0.43%
University of Notre Dame	–3,557	–36.73%	7,114	58.05%	7.36%	0.58%
Dartmouth College	–1,789	–36.09%	3,578	56.48%	6.43%	0.84%
Medical College of Wisconsin	–287	–32.77%	574	48.74%	10.98%	0.98%
Baylor College of Medicine	–351	–30.97%	702	44.86%	6.35%	0.84%
Washington and Lee University	–469	–30.32%	938	43.52%	4.13%	–0.09%
University of Richmond	–501	–21.11%	1,002	26.76%	4.22%	–0.73%
Smith College	–348	–19.72%	697	24.56%	3.88%	–1.16%
Panel B: Student above 500, and per student Asset between 500 to 600K						
Emory University	–1,109	–14.56%	2,217	17.04%	5.89%	0.37%
Claremont McKenna College	–111	–14.18%	222	16.52%	6.64%	0.93%
Icahn School of Medicine at Mount Sinai	–74	–10.90%	147	12.24%	1.94%	1.93%
University of Pennsylvania	–934	–7.65%	1,868	8.28%	11.08%	0.07%
Washington University in St Louis	–387	–5.37%	775	5.67%	5.37%	1.59%
Duke University	–302	–3.82%	604	3.97%	5.83%	0.59%
Bryn Mawr College	–22	–2.63%	45	2.70%	4.29%	0.06%
Hamilton College	–18	–1.91%	36	1.94%	5.20%	0.22%
Trinity University	–1	–0.05%	1	0.05%	3.96%	–0.11%
Panel C: Student above 500, and per student Asset between 400 to 500K						
University of Chicago	451	6.81%	–902	–6.38%	2.71%	0.89%
Berry College	89	9.20%	–178	–8.43%	4.17%	1.14%
Middlebury College	186	17.34%	–372	–14.78%	3.12%	0.04%
Northwestern University	1,515	19.06%	–3,029	–16.01%	6.65%	0.85%
Vassar College	203	20.26%	–406	–16.85%	3.71%	–0.01%
Colby College	164	21.21%	–329	–17.50%	4.25%	0.50%
Davidson College	171	23.44%	–341	–18.99%	6.18%	0.51%
Wabash College	81	23.73%	–162	–19.18%	0.15%	–0.50%

Note: The distances from the endowment threshold are calculated as the amount/number/proportion of endowment/students needed to be increased or decreased in order to make a college meet the tax threshold to be exempted from the tax or a college below the thresholds to be subject to the tax. The average growth rates were averaged from 2010 to 2016. All monetary amounts are reported in nominal values.

Figure A1: Distribution of Endowment Assets Per-student

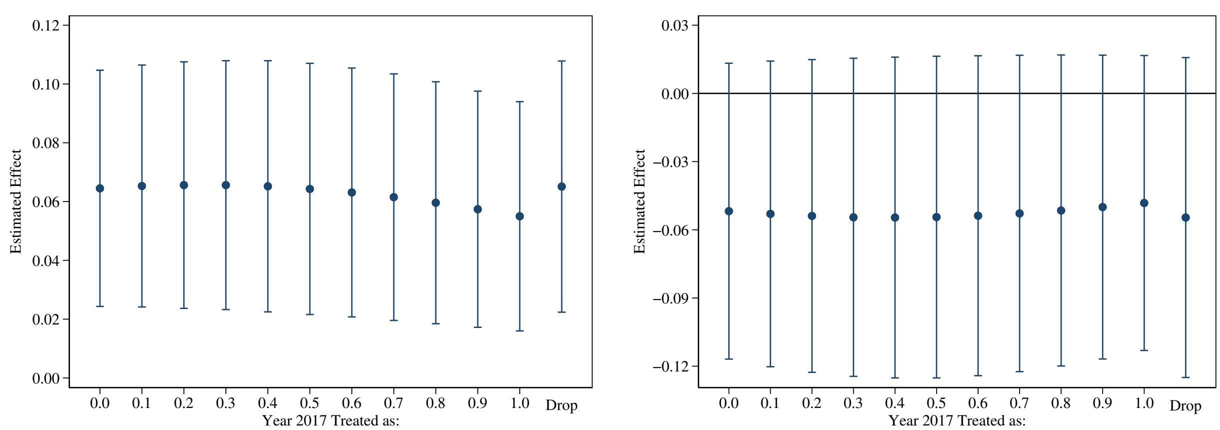
(a) Fiscal Year 2010 to 2016 | (b) Fiscal Year 2017 to 2021



Note: The samples are private nonprofit colleges that reported in IPEDS and filed Form 990 every year from 2010 to 2022. Endowment assets per student are calculated as endowment asset values divided by full-time equivalent (FTE) students (with one part-time student taken into account as one-third of full-time students). Endowment asset amounts are reported in nominal values.

Figure A2: Tax Avoidance Behaviors: Robustness Check by Definitions of Treated Period

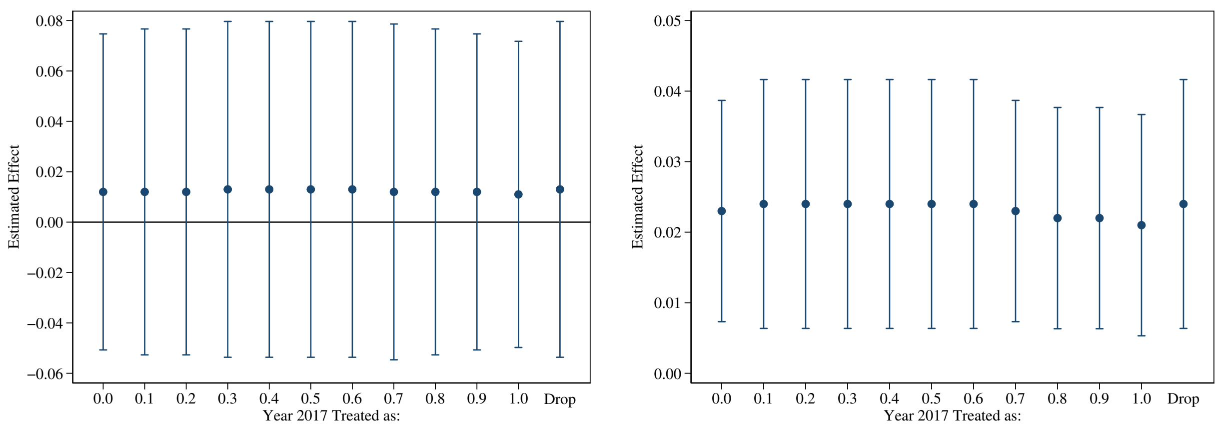
(a) Log Number of FTE Students | (b) Log Endowment Assets Per student



Note: The coefficients are estimated using equation (1). The error bars denote the 95% confidence interval. The samples are private nonprofit colleges that reported in IPEDS and filed Form 990 every year from 2010 to 2022, with a student population above 500 in 2016. FTE (full-time equivalent) is calculated as the sum of full-time and one-third of part-time students.

Figure A3: Tax Shifting Behaviors: Robustness Check by Definitions of Treated Period

(a) Log Total Expenditure (b) Log Tuition



Note: The coefficients are estimated using equation (2). The error bars denote the 95% confidence interval. The samples are private nonprofit colleges that reported in IPEDS and filed Form 990 every year from 2010 to 2022, with a student population above 500 in 2016. FTE (full-time equivalent) is calculated as the sum of full-time and one-third of part-time students.

Appendix B: Restricting Sample to Selective Colleges

B1 Empirical Design

A primary concern in the DD setting of the main analysis is the potential disparity between the treatment group (colleges taxed or near the tax threshold) and the comparison group, which typically consists of less wealthy and less selective institutions. This fundamental difference raises questions about the validity of the comparison group as a counterfactual for the treatment group.

To address this concern, I restrict the comparison group to institutions more closely resembling those in the treatment group. Beyond their wealth, most colleges subject (or potentially subject) to the NIIT are characterized by high selectivity and prestige. For instance, among the 41 colleges in our treatment group (including those taxed and those very close to the threshold), 32 are categorized as "most selective" in the Barron's Selectivity Index, three are classified as "highly competitive," and one is considered "very competitive." The remaining five are categorized as "specialized institutions." Furthermore, in the U.S. News Rankings, 32 of these colleges ranked in the top 50 (either of the ranking list of National Universities or Liberal Arts Colleges), with one ranked between 50-100 and another between 100-150.

It is reasonable to posit that colleges with similar levels of selectivity and prestige might react similarly to macroeconomic environments. These highly selective institutions typically compete with one another to attract students, and they tend to pursue similar admission strategies (Smith et al., 2018). Colleges with comparable reputations and academic rankings also tend to share similar financial metrics and management strategies (Volkwein & Sweitzer, 2006). Previous studies suggest that restricting comparisons to institutions with similar academic standing could provide a more reliable basis for analysis (Stange, 2015; Zhu et al., 2021; Bennett, 2022).

To construct more appropriate comparison groups, I link the dataset to the 2016 Barron's Selectivity Index and U.S. News rankings (for both National Universities and Liberal Arts Colleges). I created two sub-samples: one restricting to institutions in Barron's top three selectivity categories and another including those ranked in the top 100 by U.S. News in 2016.

Table B1 details the sample sizes in these sub-samples. It is important to note that while this approach restricts the comparison group to institutions more similar to the treatment group, it also excludes some treatment group institutions that are less selective and prestigious than their counterparts. This refined sample selection strategy aims to create a more comparable control group, addressing concerns about the uniqueness of the treated institutions and the potential lack of a reasonable counterfactual. By focusing on institutions with similar prestige and selectivity, we enhance the validity of our DD design, although we acknowledge the trade-off in sample size and the potential exclusion of some treated institutions.

Table B1: Number of Units in Each Sub-sample

	Number of Units	
Sub-sample	Treatment Group	Comparison Group
Tax Avoidance		
Main Results	17	752
Barrons Selectivity Index Above Very Competetive	16	268
US News' Ranking Top 100	14	108
Tax Shifting		
Main Results	24	752
Barrons Selectivity Index Above Very Competetive	20	268
US News' Ranking Top 100	19	108

B2 Empirical Results

B2.1 Tax Avoidance

Table B2 replicates the main results of colleges' manipulation behaviors related to student enrollment using our alternative, more selective samples. The findings suggest that colleges around the cutoff increased their FTE enrollment by 5.4% to 6.6%, closely aligning with our main estimate of 6.4%. This consistency across sample specifications strengthens our confidence in the robustness of these results. The decomposition results by enrollment status and education level also echo the main findings. Figure B1 demonstrates the dynamic effect based on the event study design, with the trajectory of the response aligning closely with the main results.

Table B2: Student Enrollment-related Tax Avoidance Behavior: Selective Colleges

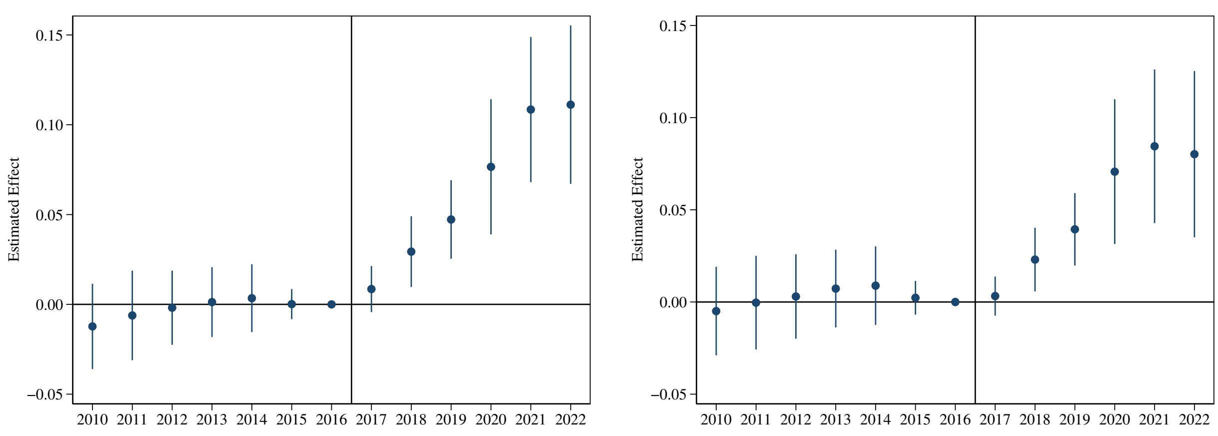
	(1)	(2)	(3)	(4)	(5)
	Log FTE	By Enrollment Status		By Student Level	
	Enrollment	Full-time	Part-time	Undergraduate	Graduate
Panel A: Barron's Rank Above Very Competetive					
$Cutoff \times Post$	0.076***	0.077***	0.024	0.068***	0.051
	(0.019)	(0.020)	(0.124)	(0.024)	(0.187)
Observations	3,640	3,640	3,640	3,640	3,640
Baseline Mean (Thousand)	7.272	6.955	0.950	4.010	3.262
Panel B: US News' Ranking Top 100					
$Cutoff \times Post$	0.057***	0.057***	0.064	0.045*	0.119
	(0.020)	(0.020)	(0.136)	(0.026)	(0.212)
Observations	1,560	1,560	1,560	1,560	1,560
Baseline Mean (Thousand)	7.988	7.630	1.072	4.274	3.714

Note: The coefficients are estimated using equation (1). Standard errors clustered at the institution level in parentheses. The outcomes are log students enrollment. The number of full-time equivalent (FTE) students is defined as the sum of full-time and one-third of part-time students. Panel A restricts the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Panel B restricts the sample to those with US News Ranking among the top 100 in 2016. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Figure B1: Tax Avoidance Behavior: Log Number of FTE Students

(a) Barron's Index Above Very Competetive

(b) US News' Ranking Top 100



Note: The coefficients are estimated using the event study version of equation (1). The error bars denote the 95% confidence interval. Samples are private non-profit colleges that reported in IPEDS and filed Form 990 yearly from 2010 to 2022, with a student population above 500 in 2016. Figure B1a restricts the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Figure B1b restricts the sample to those with US News Ranking among the top 100 in 2016.

Table B3 examines endowment asset manipulation. Consistent with the main results, I find a null response in total endowment and across various asset categories. This consistency suggests our results are not driven by differences between highly and less selective institutions. Figure B2 illustrates the event study analysis, showing temporal patterns of endowment responses mirror our main analysis, reinforcing the robustness of our results across institutional profiles.

Table B3: Endowment and Asset-related Tax Avoidance Behavior: Selective Colleges

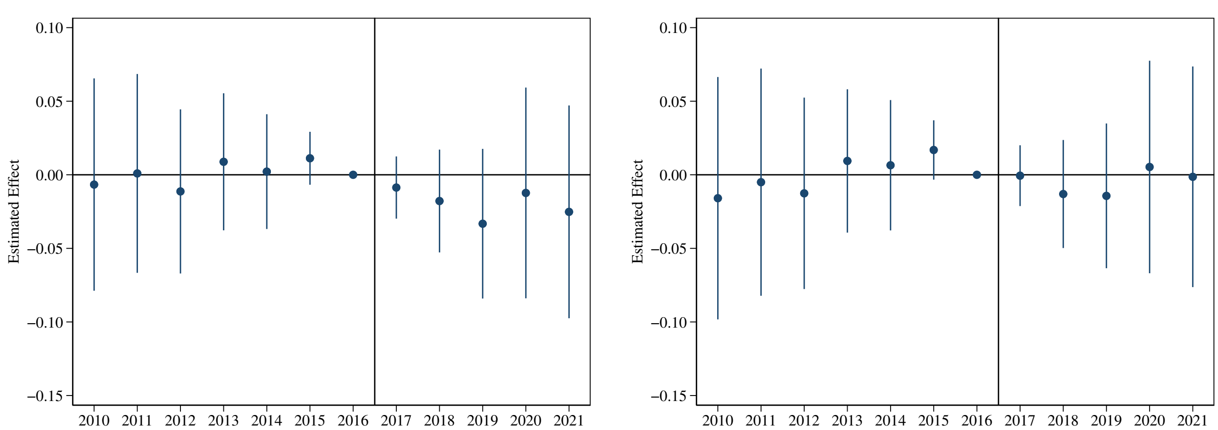
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Endowment		By Restricted Status		By Category			
	Total	Per-student	Non-restricted	Restricted	Capital	Investment	Others	Liability
Panel A: Barron's Rank Above Very Competetive								
$Cutoff \times Post$	–0.006	–0.095*	–0.041	0.047	0.071	0.024	–0.934	0.055
	(0.056)	(0.049)	(0.220)	(0.047)	(0.068)	(0.052)	(1.261)	(0.080)
Observations	3,360	3,360	3,360	3,360	3,360	3,360	3,360	3,360
Baseline Mean (Thousand)	3,637	0.481	2,338	2,480	2,865	4,439	13	1,883
Panel B: US News' Ranking Top 100								
$Cutoff \times Post$	–0.011	–0.088	–0.126	0.040	0.074	0.016	–1.101	0.030
	(0.062)	(0.054)	(0.245)	(0.049)	(0.077)	(0.059)	(1.446)	(0.091)
Observations	1,440	1,440	1,440	1,440	1,440	1,440	1,440	1,440
Baseline Mean (Thousand)	4,002	0.481	2,625	2,710	3,220	4,919	15	2,131

Note: The coefficients are estimated using equation (1). Standard errors clustered at the institution level in parentheses. The outcomes are log endowment assets. All dollars are adjusted by CPI and denoted in 2010 real dollars. Panel A restricts the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Panel B restricts the sample to those with US News Ranking among the top 100 in 2016.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Figure B2: Tax Avoidance Behavior: Log Endowmenr Per Student

(a) Barron's Index Above Very Competetive

(b) US News' Ranking Top 100



Note: The coefficients are estimated using the event study version of equation (1). The error bars denote the 95% confidence interval. Samples are private non-profit colleges that reported in IPEDS and filed Form 990 yearly from 2010 to 2022, with a student population above 500 in 2016. Figure B2a restricts the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Figure B2b restricts the sample to those with US News Ranking among the top 100 in 2016.

B2.2 Tax Shifting

Table B4 presents a similar analysis focusing on tax-shifting behaviors. The results show a null effect on total spending and in most spending categories, consistent with our main findings. The only exception is the estimate of total spending on institutional grants. While the main result shows no significant impact on institutional grants, the subsample focusing on selective colleges demonstrates a 9% to 41% increase in grant spending ($p < 0.1$). However, it is important to note that due to the smaller sample size, these estimates are less precise. Figure B3 illustrates the dynamic effects based on the event study design. The Barron's Index sample shows a pattern very similar to the main findings, while the US News sample demonstrates a minor, non-significant negative trend in total spending after policy adoption.

Table B4: Expenditure-related Tax Shifting Behavior: Selective Colleges

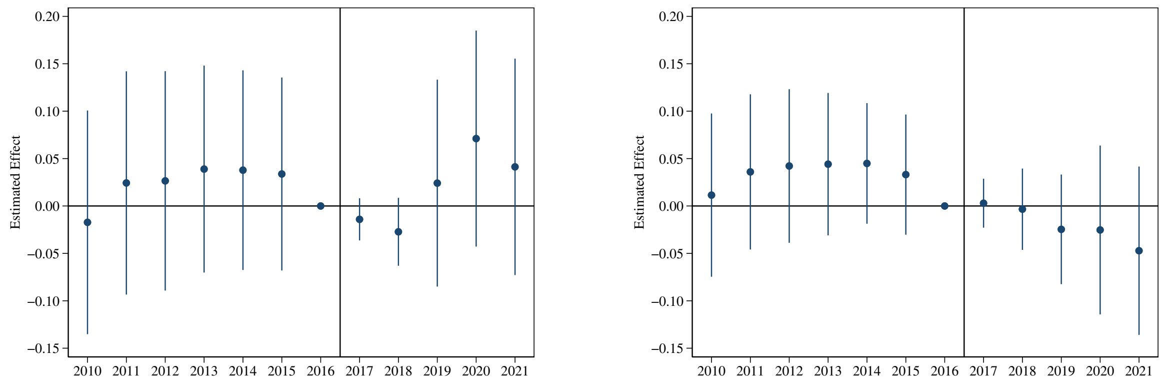
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log Expenditure						
	Total	Instruction	Research	Public Service	Institution Support	Auxiliary Facilities	Institution Grant
Panel A: Barron's Rank Above Very Competetive							
$Treat \times Post$	0.005	–0.004	0.096	0.129	–0.053	0.017	0.418*
	(0.036)	(0.038)	(0.114)	(0.140)	(0.054)	(0.050)	(0.232)
Observations	3,324	3,324	3,324	3,324	3,324	3,324	3,324
Baseline Mean (Million)	1,614	552	222	8	134	490	143
Panel B: US News' Ranking Top 100							
$Treat \times Post$	–0.055	–0.028	0.142	0.089	–0.167*	–0.107	0.093*
	(0.050)	(0.058)	(0.218)	(0.306)	(0.088)	(0.067)	(0.052)
Observations	1,380	1,380	1,380	1,380	1,380	1,380	1,380
Baseline Mean (Million)	1,731	591	239	9	143	526	151

Note: The coefficients are estimated using equation (2). Standard errors clustered at the institution level in parentheses. The outcomes are the log expenditure by spending category. Column (1) is the total expenditure. Column (2) is the sum of instructional and academic support expenditures. Column (3) is the sum of research and independent operation expenditure. Column (4) is the public service expenditure. Column (5) is the institutional support expenditure, which includes spending on operational support, administrative services, and management. Column (6) is the sum of auxiliary facilities, hospital, and student service expenditure. Column (7) is the net institutional grant aid to students, including scholarships and fellowships. All dollars are adjusted by CPI and denoted in 2010 real dollars. Samples are private non-profit colleges that reported in IPEDS and filed Form 990 yearly from 2010 to 2022, with a student population above 500 in 2016. All Panels exclude colleges with endowment assets per student between $400,000 and 600,000 in 2016 (i.e., only include the donut sample). Panel A restricts the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Panel B restricts the sample to those with US News Ranking among the top 100 in 2016. The observation period is from 2010 to 2021.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Figure B3: Tax Shifting Behavior: Log Total Expenditure

(a) Barron's Index Above Very Competetive

(b) US News' Ranking Top 100



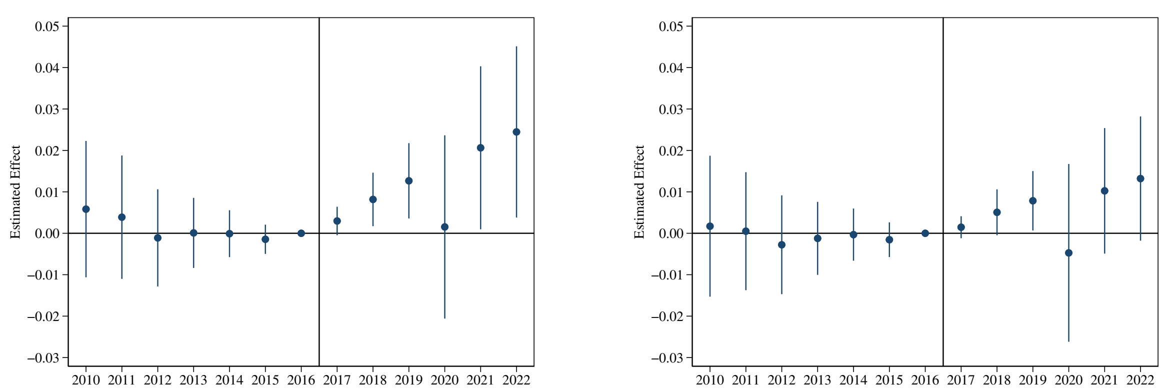
Note: The coefficients are estimated using the event study version of equation (2). The error bars denote the 95% confidence interval. Samples are private non-profit colleges that reported in IPEDS and filed Form 990 yearly from 2010 to 2022, with a student population above 500 in 2016. Figure B3a restricts the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Figure B3b restricts the sample to those with US News Ranking among the top 100 in 2016.

Table B5 investigates responses in tuition and charges. The results indicate that taxed colleges increase listed undergraduate tuition by 1.8% to 2.7%, close to the main estimate of 2.4%. Figure B4 demonstrates the dynamic effects based on the event study design. Both samples show similar trends, though the US News sample exhibits a smaller magnitude of effect. The response in room & board charges shows some variation across samples. Results restricted to institutions with higher Barron's Selectivity Index demonstrate an increase in charges of 4.4% (close to the main estimate of 4.2%), while results based on top-ranking colleges show an insignificant 1.8% response.

Figure B4: Tax Shifting Behavior: Log Listed Undergraduate Tuition

(a) Barron's Index Above Very Competetive

(b) US News' Ranking Top 100



Note: The coefficients are estimated using the event study version of equation (1). The error bars denote the 95% confidence interval. Samples are private non-profit colleges that reported in IPEDS and filed Form 990 yearly from 2010 to 2022, with a student population above 500 in 2016. FigureB4a restricts the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Figure B4b restricts the sample to those with US News Ranking among the top 100 in 2016.

Table B5: Enrollment, Tuition, and Charge-related Tax Shifting Behavior: Selective Colleges

	(1)	(2)	(3)	(4)	(5)	(6)
	Log FTE	Log Listed Price			Log Total Revenue	
	Enrollment	Undergrad Tuition	Graduate Tuition	Room & Board	Tuition	Auxiliary
Panel A: Barron's Rank Above Very Competetive						
$Treat \times Post$	0.015	0.027*	0.007	0.044***	0.109**	0.014
	(0.027)	(0.014)	(0.027)	(0.015)	(0.047)	(0.064)
Observations	3,601	3,601	3,601	3,601	3,324	3,324
Baseline Mean (Thousand)	6.917	43.415	28.498	12.995	187,940	71,791
Panel B: US News' Ranking Top 100						
$Treat \times Post$	0.009	0.018*	0.023	0.018	0.042	–0.010
	(0.025)	(0.010)	(0.030)	(0.014)	(0.050)	(0.078)
Observations	1,495	1,495	1,495	1,495	1,380	1,380
Baseline Mean (Thousand)	7.321	43.915	27.851	12.990	200,481	76,417

Note: The coefficients are estimated using equation (2). Standard errors clustered at the institution level in parentheses. The outcomes are the log enrollment, price, and revenue. All dollars are adjusted by CPI and denoted in 2010 real dollars. Samples are private non-profit colleges that reported in IPEDS and filed Form 990 yearly from 2010 to 2022, with a student population above 500 in 2016. All Panels exclude colleges with endowment assets per student between $400,000 and 600,000 in 2016 (i.e., only include the donut sample). Panel A restricts the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Panel B restricts the sample to those with US News Ranking among the top 100 in 2016. The observation period is from 2010 to 2022 for columns (1) to (4) and 2010 to 2021 for columns (5) and (6).
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

B2.3 Enrollment Composition

Table B6 explores the effects on student enrollment by race/ethnicity. In general, the results align well with the main findings. Tax avoidance behaviors lead to an increase in student enrollment across almost all racial/ethnic groups (with the exception of Black students in the US News subsample). Conversely, tax-shifting behaviors result in a significant decrease in Hispanic student enrollment. The US News sample additionally identifies a significant negative effect on Black student enrollment. These results from more selective subsamples largely corroborate our main findings, suggesting that the observed effects on enrollment composition are consistent across different institutional profiles.

Table B6: Student Enrollment-related Tax Avoidance and Shifting Behavior by Race/Ethnicity

	(1)	(2)	(3)	(4)	(5)	(6)
	Log FTE Enrollment					
	White	Black	Hispanic	Asian	Other Minority	NRA
Panel A: Tax Avoidance, Barron's Rank Above Very Competetive						
$Cutoff \times Post$	0.085***	0.019	0.106**	0.069*	0.212***	0.018
	(0.024)	(0.033)	(0.043)	(0.038)	(0.051)	(0.056)
Observations	3,900	3,900	3,900	3,900	3,900	3,900
Baseline Mean (Thousand)	3.447	0.433	0.572	0.950	0.259	1.318
Panel B: Tax Avoidance, US News' Ranking Top 100						
$Cutoff \times Post$	0.072***	–0.038	0.091***	0.046	0.184***	0.035
	(0.022)	(0.030)	(0.029)	(0.032)	(0.054)	(0.045)
Observations	1,807	1,807	1,807	1,807	1,807	1,807
Baseline Mean (Thousand)	3.723	0.480	0.607	1.074	0.283	1.495
Panel C: Tax Shifting, Barron's Rank Above Very Competetive						
$Treat \times Post$	0.002	–0.037	–0.105***	0.022	–0.018	0.111**
	(0.022)	(0.030)	(0.039)	(0.035)	(0.046)	(0.051)
Observations	3,692	3,692	3,692	3,692	3,692	3,692
Baseline Mean (Thousand)	3.135	0.387	0.592	0.956	0.304	1.336
Panel D: Tax Shifting, US News' Ranking Top 100						
$Treat \times Post$	–0.012	–0.078***	–0.124***	0.028	0.011	0.042
	(0.020)	(0.026)	(0.026)	(0.028)	(0.048)	(0.040)
Observations	1,625	1,625	1,625	1,625	1,625	1,625
Baseline Mean (Thousand)	3.315	0.412	0.615	1.017	0.319	1.426

Note: The coefficients in Panel A and B are estimated using equation (1). The coefficients in Panel C and D are estimated using equation (2). Standard errors clustered at the institution level in parentheses. The outcomes are log full-time equivalent (FTE) students by race/ethnicity. Other minorities include Native Hawaiian and Pacific Islander (NHPI), American Indians and Alaska Natives (AIAN), and two or more races. NRA stands for non-resident alien. Samples are private non-profit colleges that reported in IPEDS and filed Form 990 yearly from 2010 to 2022, with a student population above 500 in 2016. Panels C and D exclude colleges with endowment assets per student between $400,000 and 600,000 in 2016 (i.e., only include the donut sample). Panel A and B restrict the sample to those with Barron's Rank as most competitive, highly competitive, or very competitive. Panels C and D restrict the sample to those with US News Ranking among the top 100 in 2016. The observation period is from 2010 to 2022.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Appendix C: Triple-Difference Design

C1 Empirical Strategy

In the main analysis, I use the DD framework to estimate colleges' tax avoidance and shifting behaviors. In the tax avoidance analysis, the treatment group consists of colleges near the asset threshold of the NIIT, while the comparison groups include those far away from the threshold. In the tax shifting analysis, I compare colleges subjected to the tax (treatment group) with those that meet the student threshold but not the asset threshold (the comparison group). However, in both settings, given the substantial difference between treatment and comparison groups, concern exists about whether they would have shared the same trend in the outcome variables. Despite the event study analysis demonstrating a parallel pre-treatment trend (at least conditional on the fixed effect), the concern of the DD setting still remains.

Hence, this study further applies a triple-difference (DDD) framework to test the robustness of the results. In the main analysis, I only included colleges that met the student threshold and separated the samples into treatment and comparison groups depending on the distance to the assets threshold. In the DDD design, I further introduce those colleges that do not meet the student threshold as additional comparison groups. The setup slightly differs between the tax avoidance and shifting analysis.

C1.1 Tax Avoidance

In the tax avoidance analysis, I compare the colleges around the assets threshold and those far away between those meeting the student threshold and those not. In other words, I compare two differences: (1) the difference between the cutoff sample (with endowment assets per student within $400,000 to $600,000 in 2016) and the non-cutoff sample within large colleges (with student enrollment greater than 500 in 2016); (2) the same difference but within small colleges (with student enrollment less than 500 in 2016). And then, I track the change in the gaps between these two differences across time. Specifically, I estimate the following equation:

$$Y_{it} = \alpha_0 + \beta_1 Large_i \times Cutoff_i \times Post_t \tag{C1}$$
$$+ \theta_i + Large_i \times \delta_t + Cutoff_i \times \zeta_t + Above_i \times \phi_t + \varepsilon_{it}$$

Where $Large_i$ is a dummy variable indicating that the colleges had a student population above 500 in 2016. $Cutoff_i$ is a dummy variable indicating that the colleges had endowment assets per student within \$400,000 and \$600,000 in 2016. The equation includes the student population by year fixed effect ($Large_i \times \delta_t$), which accounts for the potential difference in trends between large and small colleges. Similarly, the inclusion of the distance to the cutoff status by year fixed effect ($Cutoff_i \times \zeta_t$) accounts for the potential difference in trends between those colleges that have similar levels of wealth and those not. θ_i is the institution fixed effect, which absorbs the interaction term of $Large_i \times Cutoff_i$. These three terms stand for the full interactions to establish the DDD setting. Similar to the equation (1), the specification includes the above-cutoff-status-by-year fixed effects ($Above_i \times \phi_t$) to account for potential differences in trends between those subject and those not subject to the tax. The key parameter is β_1, which indicates the behavioral response of colleges that have the motivation of tax avoidance.

The empirical assumption of the DDD is that the difference in outcomes between "large, around assets cutoff" and "large, not around assets cutoff" colleges would have followed the same trend as the difference between "small, around assets cutoff" and "small, not around assets cutoff" colleges in the absence of the policy. This assumption might be valid as the primary factors determining colleges' enrollment and finance metrics would be their student body and available resources.

C1.2 Tax Shifting

In the tax shifting analysis, I separate colleges into four groups by both the student and assets threshold. Colleges meeting the student threshold (with student enrollment greater than 500 in 2016) are categorized as large and small otherwise. Colleges meeting the asset threshold (with endowment assets per student above \$500,000 in 2016) are categorized as wealthy and non-wealthy otherwise. As demonstrated in Figure 1a, this categorization groups colleges into four quadrants, with the upper right corner denoting the treatment group.

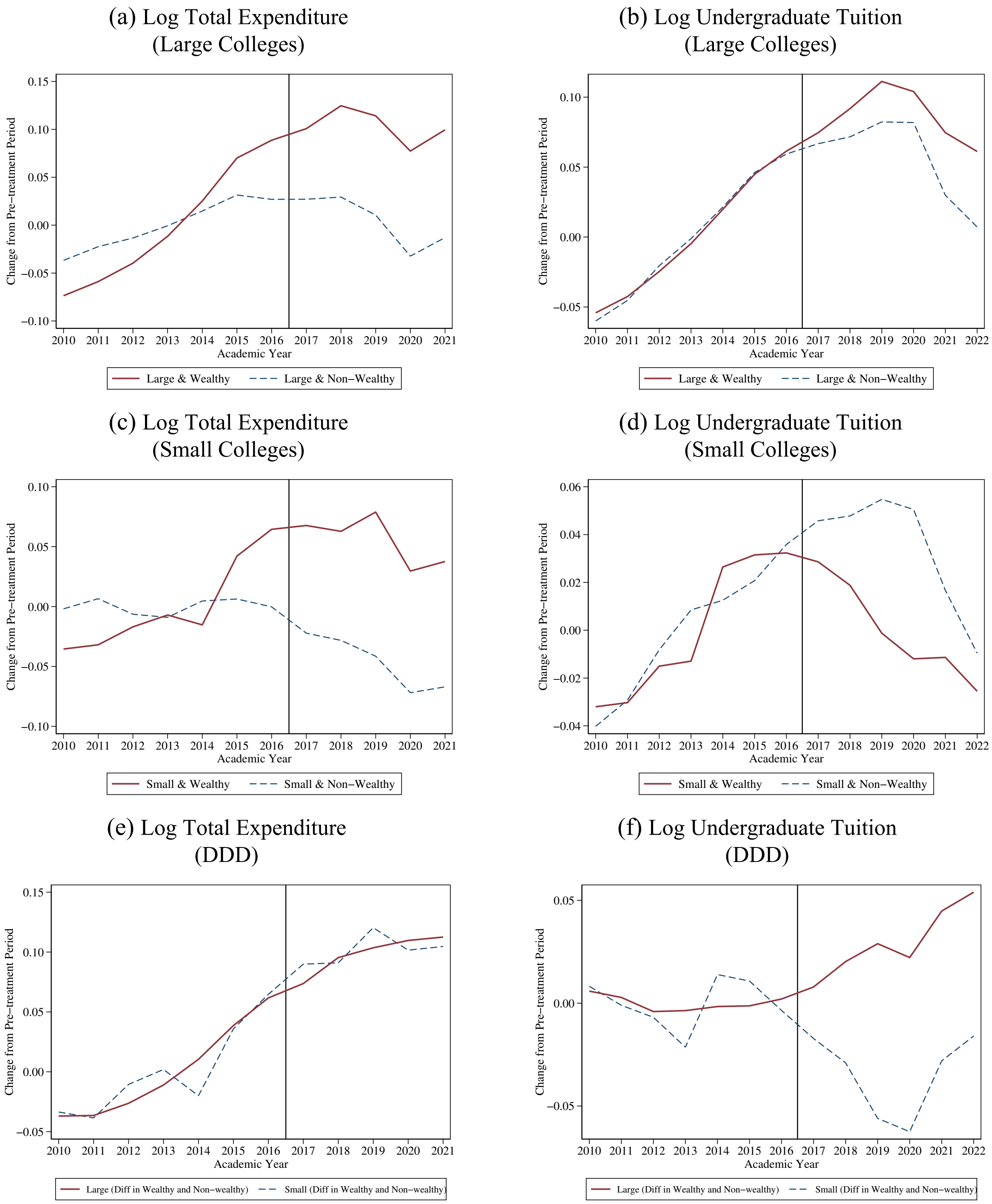
The basic idea of the DDD setting is to compare the changes in the gap between large wealthy and large non-wealthy colleges as well as the gap between small wealthy and small non-wealthy colleges. This analysis consists of all colleges (including those that unmet the student threshold) but still excludes those around the cutoff to prevent confounding from tax avoidance behaviors. Specifically, I estimate the following equation:

$$Y_{it} = \alpha_0 + \beta_1 Large_i \times Wealthy_i \times Post_t + \theta_i + Large_i \times \delta_t + Wealthy_i \times \zeta_t + \varepsilon_{it} \quad \text{(C2)}$$

Where $Large_i$ is a dummy variable indicating that the colleges had a student population above 500 in 2016. $Wealthy_i$ is a dummy variable indicating that the colleges had endowment assets per student above \$500,000 in 2016. The equation includes the student population by year fixed effect ($Large_i \times \delta_t$), which accounts for the potential difference in trends between large and small colleges. Similarly, the inclusion of asset size by year fixed effect ($Wealthy_i \times \zeta_t$) accounts for the potential difference in trends between wealthy and non-wealthy colleges. θ_i is the institution fixed effect, which absorbs the interaction term of $Large_i \times Wealthy_i$. These three terms stand for the full interactions to establish the DDD setting. The key parameter is β_1, which indicates the impact of policy on the colleges subject to the NIIT.

The empirical assumption of the DDD setting is that the difference in outcomes between “large, wealthy” and “large, non-wealthy” colleges would have followed the same trend as the difference between “small, wealthy” and “small, non-wealthy” colleges in the absence of the policy. In other words, the DDD design assumes that the gap between wealthy and non-wealthy colleges would be the same between colleges with various student sizes. This assumption might be valid as the primary factors determining colleges’ finance metrics would be their service population and available resources. This paper further evaluates the assumption by examining the pre-treatment parallel trend. Specifically, while “large, wealthy colleges” (treated group) hold a faster growth rate in expenditure than the “large, non-wealthy colleges” (see Figure C1a), the same pattern appears in the comparison between “small, wealthy colleges” versus “small, non-wealthy colleges” (see Figure C1c). Figure C1e compares the gap in two paired comparisons and shows the same trend over time.

Figure C1: Tax Shifting: Trend in Total Expenditure and Tuition



Note: The samples are private nonprofit colleges that reported in IPEDS and filed Form 990 every year from 2010 to 2022 and exclude colleges with endowment assets per student between \$400,000 and 600,000 in 2016 (i.e., only include the donut sample). The horizontal axis denotes the year (using the start year of the academic/fiscal year). The vertical axis denotes the percent change in the outcome variable from the pre-treatment period. The vertical line denotes the year of policy implementation. Large (small) colleges are colleges with more (less) than 500 students in 2016. Wealthy (non-wealthy) colleges are colleges with more (less) than \$500,000 endowment assets per student (in nominal values) in 2016.

This paper employs DD in the primary setting while using DDD as a robustness check. The choice of the preferred specification involves a trade-off between bias and precision. While the DDD framework is better suited to correct the bias of comparing colleges with different asset levels, it necessitates the introduction of a comparison group of small but wealthy colleges. Most of these colleges are arts or medical schools. Due to their small student population and significant assets, they typically experience frequent and substantial fluctuations in spending. This setting, therefore, introduces more noise to the estimation and leads to larger standard errors.

C2 Empirical Results

C2.1 Tax Avoidance

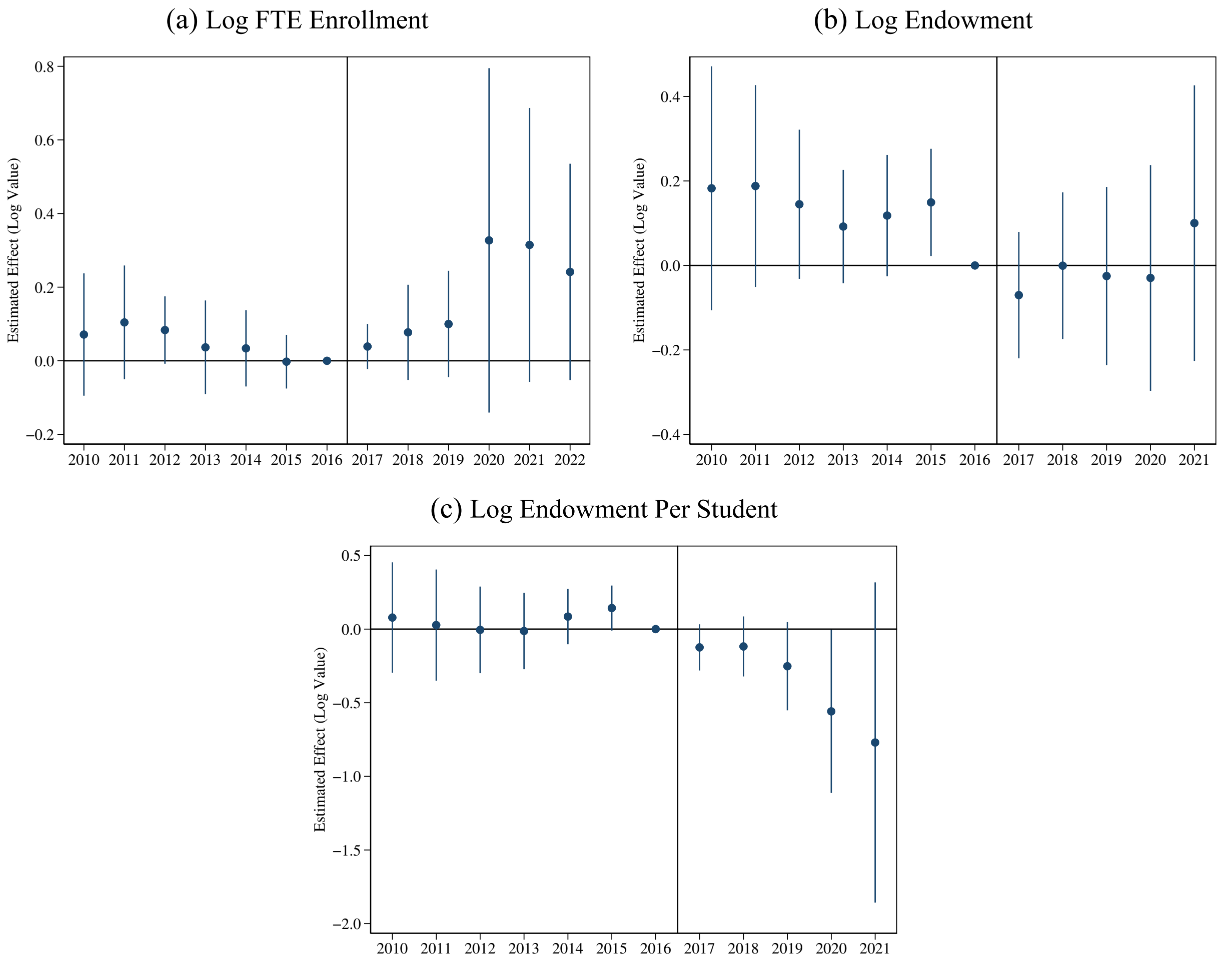
The DDD results of the student enrollment-related tax avoidance align with the main findings, though with larger standard errors. Table C1 demonstrates that colleges around the cutoff increase their FTE enrollment by 16% after the policy implementation, despite the estimate being non-significant. The event study results in Figure C2a show a good pre-treatment common trend and a clear pattern of increase in enrollment among the "large and around cutoff" group, despite all estimates being non-significant. The noisier estimates are likely due to fluctuations in student enrollment among smaller colleges. Despite this limitation, the pattern in student enrollment among affected colleges is still evident and aligns with the main findings.

Table C1: Student Enrollment-related Tax Avoidance Behavior: DDD Setting

	(1)	(2)	(3)	(4)	(5)
	Log FTE	By Enrollment Status		By Student Level	
	Enrollment	Full-time	Part-time	Undergraduate	Graduate
$Large \times Cutoff \times Post$	0.181	0.198	0.095	0.066	0.213
	(0.134)	(0.162)	(0.139)	(0.100)	(0.248)
Observations	11,661	11,661	11,661	11,661	11,661
Baseline Mean (Thousand)	4.928	4.715	0.639	2.678	2.250

Note: The coefficients are estimated using equation (C1). Standard errors clustered at the institution level in parentheses. The outcomes are log students enrollment. The number of full-time equivalent (FTE) students is defined as the sum of full-time and one-third of part-time students.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Figure C2: Event Study Estimates: Avoidance Behavior



Note: The coefficients are estimated using the event study version of equation (C2). The error bars denote the 95% confidence interval. The samples are private nonprofit colleges that reported in IPEDS and filed Form 990 every year from 2010 to 2022, and exclude colleges with endowment assets per student between \$400,000 and 600,000 in 2016 (i.e., only include the donut sample).

Table C2 explores the manipulation behaviors on endowment assets. The results demonstrate a 12% non-significant drop in total endowment for colleges with motivation for tax avoidance. Despite the non-trivial point estimate, the event study result in Figure C2b shows no clear pattern of a drop in asset values after the policy implementation. The overall findings still align with the main results.

Due to the increase in enrollment and unchanged endowment assets, the results show a drop in endowment assets per student (see Column (2) in Table C2 and Figure C2c). Overall, the tax avoidance findings from the DDD setting corroborate the main findings.

Table C2: Tax Avoidance Behavior on Student Enrollment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Endowment		By Restricted Status		By Category			
	Total	Per-student	Non-restricted	Restricted	Capital	Investment	Others	Liability
$Large \times Cutoff \times Post$	–0.136	–0.536**	–0.041	0.126**	0.050	0.096	–0.873	–0.080
	(0.132)	(0.249)	(0.234)	(0.063)	(0.094)	(0.076)	(0.937)	(0.186)
Observations	10,764	10,764	10,764	10,764	10,764	10,764	10,764	10,764
Baseline Mean (Thousand)	2,249	0.442	1,432	1,541	1,811	2,731	8	1,196

Note: The coefficients are estimated using equation (C1). Standard errors clustered at the institution level in parentheses. The outcomes are log endowment assets. All dollars are adjusted by CPI and denoted in 2010 real dollars. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

C2.2 Tax Shifting

The DDD results of the tax shifting response on total expenditure are quite similar to the DD estimations. Table C3 demonstrates that taxed colleges underwent an insignificant 0.2% increase in their total expenditure after the policy intervention (see Column (1)). There are also no negative responses for any of the spending categories.

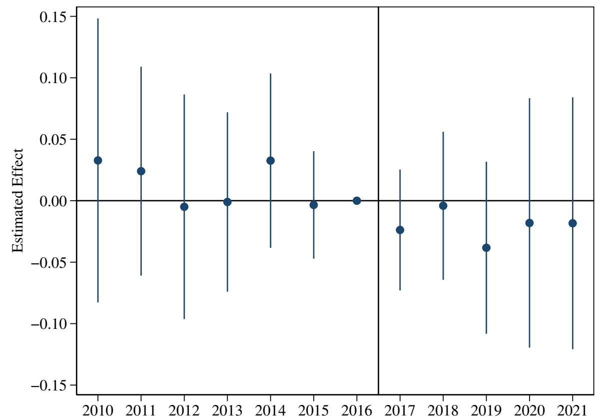
Table C3: Expenditure-related Tax Shifting Behavior: DDD Setting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log Expenditure						
	Total	Instruction	Research	Public Service	Institution Support	Auxiliary Facilities	Institution Grant
$Large \times Wealthy \times Post$	0.002	0.025	0.029	0.140**	0.102	0.042	0.249
	(0.043)	(0.043)	(0.054)	(0.059)	(0.108)	(0.047)	(0.411)
Observations	11,004	11,004	11,004	11,004	11,004	11,004	11,004
Baseline Mean (Million)	1,524	478	222	28	121	459	123

Note: The coefficients are estimated using equation (3). Standard errors clustered at the institution level in parentheses. The outcomes are the log expenditure by spending category. Column (1) is the total expenditure. Column (2) is the sum of instructional and academic support expenditures. Column (3) is the sum of research and independent operation expenditure. Column (4) is the public service expenditure. Column (5) is the institutional support expenditure, which includes spending on operational support, administrative services, and management. Column (6) is the sum of auxiliary facilities, hospital, and student service expenditure. Column (7) is the net institutional grant aid to students, including scholarships and fellowships. All dollars are adjusted by CPI and denoted in 2010 real dollars. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

The event-study estimation reassures the findings. Figure C3 demonstrates non-significant estimates for all the pre-intervention periods, showing a good common trend. The results also suggest a null effect on spending change after the policy intervention.

Figure C3: Event Study Estimates: Tax Shifting Behavior on Total Expenditure



Note: The coefficients are estimated using the event study version of equation (C2). The error bars denote the 95% confidence interval. The samples are private nonprofit colleges that reported in IPEDS and filed Form 990 every year from 2010 to 2022, and exclude colleges with endowment assets per student between $400,000 and 600,000 in 2016 (i.e., only include the donut sample).

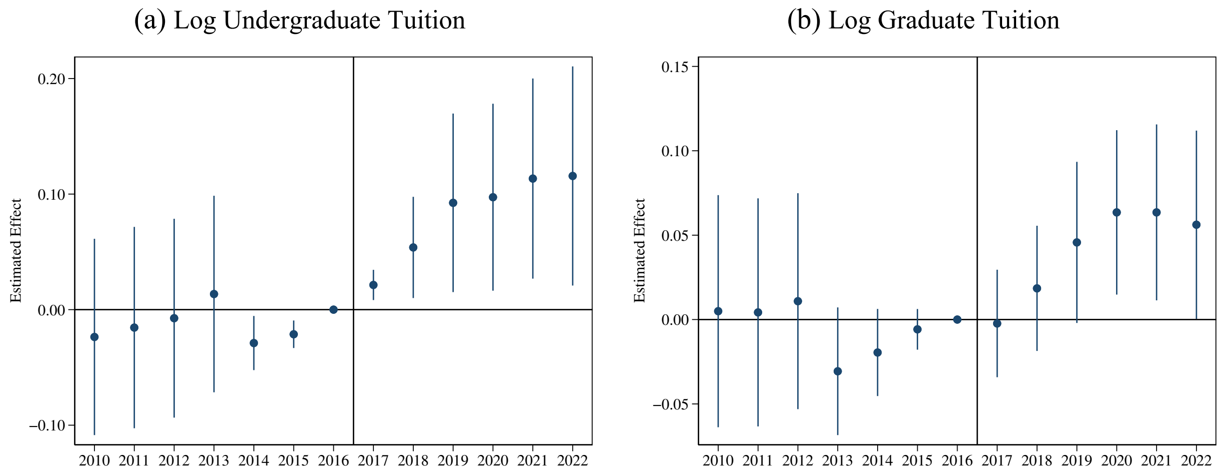
The results on tuition hikes align with the DD results but with larger estimates. Table C4 finds that taxed colleges underwent a 10% increase in undergraduate tuition ($p < 0.01$, see Column (2)), 5% increase in graduate tuition ($p < 0.1$, see Column (3)), and 6% increase in room and board charge ($p < 0.01$, see Column (4)). Despite the larger magnitude, the 95% confidence intervals overlap with the DD estimates. The event-study estimates (see Figure C4), confirm the parallel trend in the pre-intervention period and show that the tuition has gradually increased over time.

Table C4: Enrollment, Tuition, and Charge-related Tax Shifting Behavior: DDD Setting

	(1)	(2)	(3)	(4)	(5)	(6)
	Log FTE	Log Listed Price			Log Total Revenue	
	Enrollment	Undergrad Tuition	Graduate Tuition	Room & Board	Tuition	Auxiliary
$Large \times Wealthy \times Post$	–0.084	0.100***	0.052*	0.059***	0.214	–0.138
	(0.079)	(0.033)	(0.028)	(0.017)	(0.212)	(0.135)
Observations	11,004	11,004	11,004	11,004	11,004	11,004
Baseline Mean (Thousand)	6.037	39.033	28.449	11.451	162,878	61,246

Note: The coefficients are estimated using equation (3). Standard errors clustered at the institution level in parentheses. The outcomes are the log expenditure by spending category. Column (1) is the total expenditure. Column (2) is the sum of instructional and academic support expenditures. Column (3) is the sum of research and independent operation expenditure. Column (4) is the public service expenditure. Column (5) is the institutional support expenditure, which includes spending on operational support, administrative services, and management. Column (6) is the sum of auxiliary facilities, hospital, and student service expenditure. Column (7) is the net institutional grant aid to students, including scholarships and fellowships. All dollars are adjusted by CPI and denoted in 2010 real dollars.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Figure C4: Event Study Estimates: Tax Shifting Behavior



Note: The coefficients are estimated using the event study version of equation (C2). The error bars denote the 95% confidence interval. The samples are private nonprofit colleges that reported in IPEDS and filed Form 990 every year from 2010 to 2022, and exclude colleges with endowment assets per student between $400,000 and 600,000 in 2016 (i.e., only include the donut sample).

Figure C1 provides insight into the inconsistency in effect sizes between the DD and DDD models. As demonstrated in Figure C1b, colleges that are large and wealthy (subjected to the tax) show a parallel trend in tuition with colleges that are large but non-wealthy (the comparison group in the DD model) prior to the policy. However, the treatment group increased their tuition relatively more than the comparison group after the policy was effective. Despite the good pre-treatment common trend implying that large but non-wealthy colleges could serve as a good counterfactual, concerns remain about whether the common trend assumption would continue to hold true. Particularly, the pandemic might serve as a potential factor that affects the two groups differently.

This concern is backed up by evidence from the second control group from the DDD model. Figure C1d demonstrates that small but wealthy colleges and small and non-wealthy colleges also possess parallel trends prior to the policy, although these groups are more fluctuate due to their small nature. However, small but wealthy colleges show a larger drop in their tuition level during the pandemic period. One explanation could be that they are more able to use their assets to support students with a lower tuition level during hard times. The suspicion is aligned with previous studies' perspective that endowment assets could serve as the "rainy day fund" (Baum & Lee, 2019; Rosen & Sappington, 2019). In the DDD model, the response of small wealthy colleges could serve as a counterfactual for how large wealthy colleges would respond to the macro environment. Since the DDD model predicts that the treated colleges should have been able to control their tuition at a lower level as the small wealthy colleges did, the model produces a causal estimate of a larger relative increase in tuition for the treated colleges. Whether small wealthy colleges could serve as a better counterfactual for the treatment group than large non-wealthy colleges is untestable. Therefore, this paper presents the DD estimate as the lower bound while the DDD estimate as the higher bound.

Overall, the DDD estimates are generally aligned with the DD results. The evidence suggests that taxed colleges do not respond to the taxation by cutting spending but might increase tuition to shift the burden.

Appendix D: Methodology Details on Permutation Test for SCM

This paper utilize the Synthetic Control Method (SCM) to examine the treatment effect on individual institution. The conventional SCM only offer point estimates but not inference statistics. To obtain the inference statistics, this paper obtains the distribution of the estimates using a permutation test. Specifically, I perform the following steps:

Step 1: Applying SCM to placebo units:

In this step, I take each of the units in the donor pool and perform the SCM (using equation (3)). For the analysis on tax avoidance, there were 800 colleges in the donor pool; and in the tax shifting analysis, there were 752 colleges in the donor pool (see Table C1). In this permutation test, the units in the treatment group are excluded from the analysis. The practice in this step provides 800 (752) placebo estimates on each of the single units in the donor pool.

Table C1: Number of Units in Each Analysis

	Number of Units	
Analysis	Treatment Group	Donor Pool
Tax Avoidance	17	800
Tax Shifting	24	752

Step 2: Estimating placebo treatment effects:

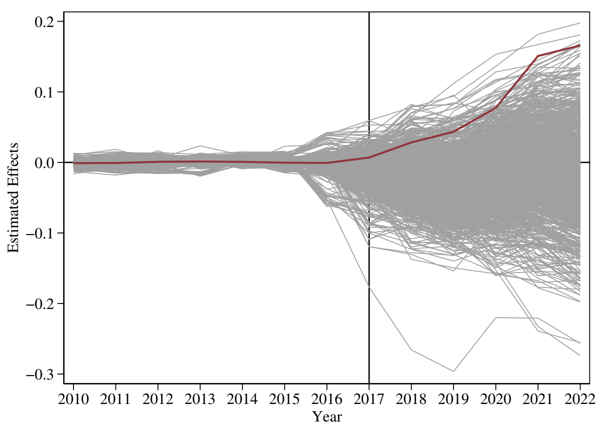
In this step, I randomly select N placebo estimates from the previous step and calculate the average treatment effect at each time period ($\overline{\beta_t}$; using equation (C1)). The number N is defined with the actual number of treated units. For example, in the tax avoidance analysis, I randomly selected 17 placebo estimates to take the average; and in the tax shifting analysis, the number would be 24. The procedure is then repeated 1,000 times, resulting in a distribution of the estimates.

$$\overline{\beta_t} = \frac{1}{N}\sum_{i=1}^{N}\beta_{it} \tag{C1}$$

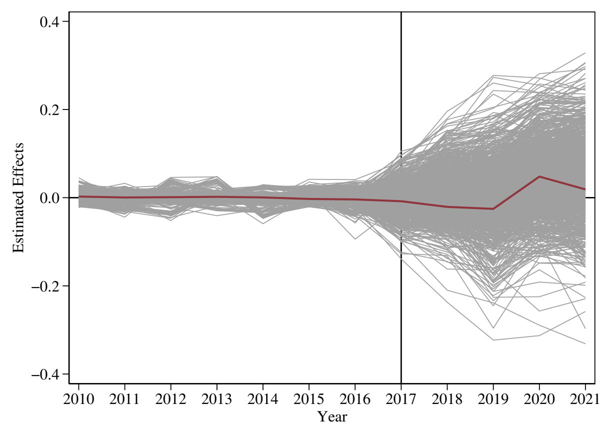
By this stage, I can already compare the actual estimates with the placebo ones to obtain the permutation p-values (for a single time period). Figure C1 demonstrates the distribution of the placebo estimates placed along with the actual estimates. These placebo estimates serve as the potential distribution of the estimated $\overline{\beta_t}$ in the absence of the policy. If the actual estimate is located at the range out of most (such as 95%) of the placebo estimates, then the estimated policy effect is likely not due to random. For the estimation of the impact of tax avoidance behavior on student enrollment, the results suggest that the actual estimate is located at the upper bound of the placebo estimates, especially in the latter year (see Figure C1a). For the estimation of the impact of tax-shifting behavior on tuition revenue, the actual estimate is also located at the upper bound of the placebo estimates (see Figure C1d).

Figure C1: SCM Permutation Test: Dynamic Treatment Effect

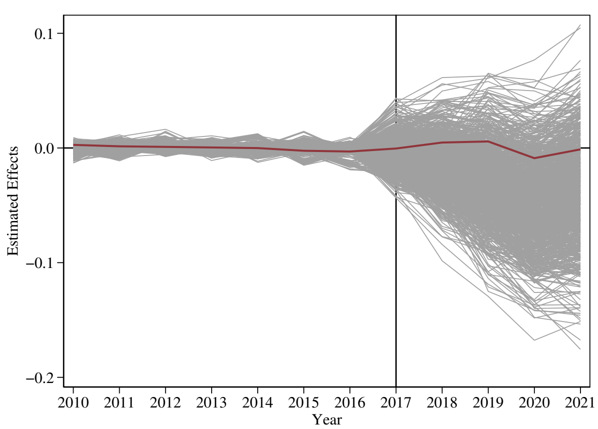
(a) Tax Avoidance:
Log Number of FTE Students



(b) Tax Avoidance:
Log Endowment Assets Per-Student



(c) Tax Shifting: Log Total Expenditure



(d) Tax Shifting: Log Tuition Revenue

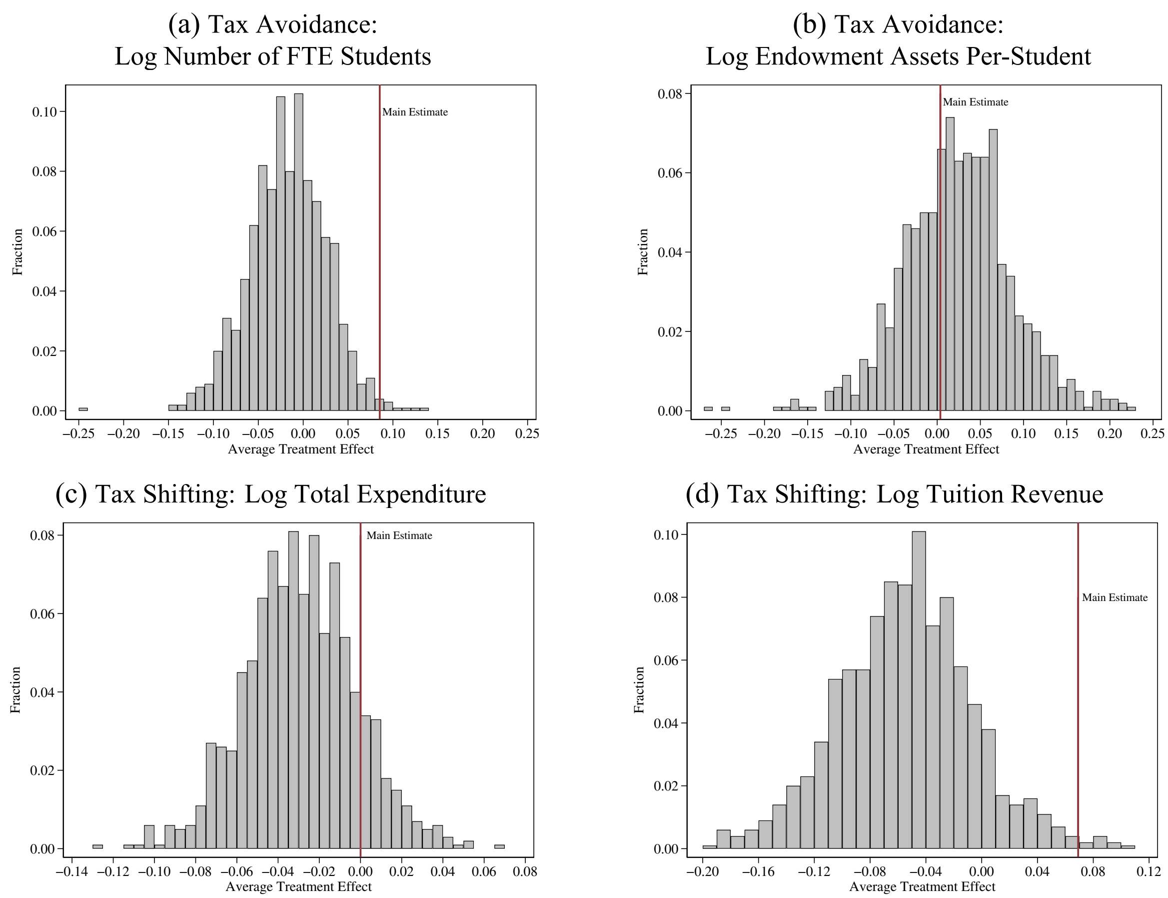
Step 3: Calculating permutation p-value for ATT:

The former step obtains the dynamic treatment effect for the placebo units. I then apply equation (C2) to compute the ATT for the entire post-treatment period.

$$ATT = \frac{1}{T - T_0 + 0.5}\left(0.5 \times \overline{\beta_{t=T_0}} + \sum_{t>T_0}^{T} \overline{\beta_t}\right) \tag{C2}$$

Figure C2 demonstrates the distribution of placebo estimates (the histogram) and the location of the actual ATT (vertical line). The permutation p-value is calculated by counting the number of placebo estimates in excess of the actual estimate. In the case of tax avoidance impact on enrollment, the permutation p-value is 0.008 as only 8 out of 1000 placebo ratio excess the actual value (see Figure C2a). Table C2 to C5 report the ATT and permutation p-value of each variable.

Figure C2: SCM Permutation Test: Average Treatment Effect



(a) Tax Avoidance: Log Number of FTE Students

(b) Tax Avoidance: Log Endowment Assets Per-Student

(c) Tax Shifting: Log Total Expenditure

(d) Tax Shifting: Log Tuition Revenue

Table C2: Enrollment-related Tax Avoidance Behavior: SCM Results

	(1)	(2)	(3)	(4)	(5)
	Log FTE	By Enrollment Status		By Student Level	
	Enrollment	Full-time	Part-time	Undergraduate	Graduate
ATT	0.085***	0.071***	–0.054	0.075*	0.033
Permutation p-value	0.008	0.004	0.694	0.057	0.144
Range	[0.029,0.182]	[–0.016,0.201]	[–0.729,0.388]	[–0.013,0.147]	[–0.191,1.095]

Note: The ATT are estimated using equation (C2). The permutation p-values are estimated using Step 3 in Appendix C. Range denotes the minimum and maximum single-institution treatment effect.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table C3: Endowment and Assets-related Tax Avoidance Behavior: SCM Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Endowment		By Restricted Status		By Category			
	Total	Per-student	Non-restricted	Restricted	Capital	Investment	Others	Liability
ATT	0.060	0.004	0.285	0.103*	0.028	0.107*	0.076	0.070**
Permutation p-value	0.121	0.647	0.161	0.060	0.117	0.075	0.599	0.046
Range	[–0.10,0.18]	[–0.13,0.16]	[–0.27,1.49]	[–0.10,0.27]	[–0.08,0.31]	[–0.05,0.46]	[–11.12,12.10]	[–0.39,0.94]

Note: The ATT are estimated using equation (C2). The permutation p-values are estimated using Step 3 in Appendix C. Range denotes the minimum and maximum single-institution treatment effect.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table C4: Expenditure-related Tax Shifting Behavior: SCM Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log Expenditure						
	Total	Instruction	Research	Public Service	Institution Support	Auxiliary Facilities	Institution Grant
ATT	0.000	0.076**	0.049	0.208*	0.023*	0.002	–0.151
Permutation p-value	0.135	0.024	0.386	0.058	0.058	0.166	0.699
Range	[–0.16,0.12]	[–0.08,0.25]	[–0.27,0.28]	[–0.15,1.10]	[–0.16,0.40]	[–0.28,0.34]	[–0.51,0.13]

Note: The ATT are estimated using equation (C2). The permutation p-values are estimated using Step 3 in Appendix C. Range denotes the minimum and maximum single-institution treatment effect.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table C5: Enrollment, Tuition, and Charge-related Tax Shifting Behavior: SCM Results

	(1)	(2)	(3)	(4)	(5)	(6)
	Log FTE	Log Listed Price			Log Total Revenue	
	Enrollment	Undergrad Tuition	Graduate Tuition	Room & Board	Tuition	Auxiliary
ATT	0.040**	0.035*	0.016	0.018***	0.069**	–0.013
Permutation p-value	0.040	0.050	0.155	0.009	0.010	0.254
Range	[–0.14,0.21]	[–0.05,0.08]	[–0.29,0.17]	[–0.14,0.18]	[–0.04,0.27]	[–0.65,0.38]

Note: The ATT are estimated using equation (C2). The permutation p-values are estimated using Step 3 in Appendix C. Range denotes the minimum and maximum single-institution treatment effect.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Step 4: Calculating permutation p-value for single unit:

To estimate the permutation p-value for single institution, I follow the approach outlined in Abadie et al. (2010) to compute the post/pre mean squared prediction error (MSPE) ratio using the following equation:

$$MSPE\,ratio = \frac{\frac{1}{T-T_0}\sum_{t>T_0}^{T}\left(\overline{\beta_t}\right)^2}{\frac{1}{T_0-1}\sum_{t<T_0}^{T_0-1}\left(\overline{\beta_t}\right)^2} \tag{C3}$$

Next, I compared the ratios of the actual estimate to the placebo estimates. The permutation p-value is calculated by counting the number of placebo post/pre-MSPE ratios in excess of the actual ratio. The level of significance of each institution is noted in the Figures 4 and 7 in the manuscript.

Appendix E: Estimation Net Benefit of Enrollment Expansion

This section estimates the net benefit derived from the enrollment expansion due to tax avoidance behavior. The estimation here is primarily based on the full-time undergraduate students as this group is the major driver of the enrollment effect. I perform the following steps to estimate the net benefits:

Step 1: Estimated the increase in college degree holders

Based on the SCM estimation, the 17 colleges around the tax threshold collectively increased their full-time undergraduate enrollment by 9,623 as of 2022. Table E1 reports the estimation for each college. Applying the degree completion rate at these colleges, this increase in enrollment could eventually result in an additional 8,799 college degree holders.[29]

Step 2: Obtained the net benefit of a college degree from prior studies

Previous studies have estimated the net personal benefit of earning a college degree to range from \$250 thousand to \$625 thousand (Hill et al., 2005; P. Taylor et al., 2011; Trostel, 2015), while the net social benefit falls between \$350 thousand and \$600 thousand (Hill et al., 2005; Edelson, 2016; Trostel, 2015). Combining the upper (lower) bounds of these estimates yields a total of \$1,225 (\$600) thousand. The estimations of individual benefits primarily hinge on the increase in earnings attributable to the degree, deducted tuition costs, and forgone earnings during college. Conversely, estimations of societal benefit primarily rely on the tax revenue accrued by the government due to increased labor earnings, net of government investment in higher education.

Step 3: Estimated the premium in return for sample college to less selective colleges

The increase in degree holders among these colleges might not be "additional." It is possible that these students could have enrolled in another college had these colleges not expanded their access. Therefore, I assume that the expansion in enrollment access essentially "moves up" stu-

[29] The average degree completion rate within 150% of normal time (i.e., 6 years) at these colleges is 88%, ranging from 65% to 97%. The estimation of degree holders is based on applying the degree completion rate in a specific college to the estimate of increased enrollment in the same college.

Table E1: Estimation of Net Benefit from Enrollment Expansion

	Barron's Ranking	Increase in FT Undergrad	Average Degree Completion Rate	Increase in Bachelor Degree	Estimate Net Benefit ($ Million)
University of Chicago	Most competitive	1,695	0.956	1,620	65.118
Emory University	Most competitive	1,481	0.900	1,333	53.572
Northwestern University	Most competitive	941	0.965	908	36.505
Washington University in St Louis	Most competitive	872	0.937	817	32.857
University of Pennsylvania	Most competitive	741	0.961	712	28.609
Duke University	Most competitive	701	0.966	677	27.234
Colby College	Most competitive	538	0.880	474	19.043
Middlebury College	Most competitive	499	0.935	467	18.781
Vassar College	Most competitive	482	0.920	443	17.803
Berry College	Very competitive	457	0.647	296	8.345
Hamilton College	Most competitive	357	0.924	330	13.259
Davidson College	Most competitive	288	0.916	264	10.593
Trinity University	Highly competitive	246	0.758	187	7.502
Claremont McKenna College	Most competitive	139	0.913	127	5.088
Wabash College	Highly competitive	119	0.753	90	3.607
Bryn Mawr College	Most competitive	67	0.826	56	2.240
Mount Sinai School of Medicine	Special	0.09	N/A†	0	0.000
Total		9,623		8,799	350

Note: The Barron's Ranking is obtained from Barron's Profiles of American Colleges, which categorizes colleges into seven categories: most competitive, highly competitive, very competitive, competitive, less competitive, noncompetitive, and special (usually art or medical schools). The increase in full-time undergraduate enrollment is measured as of 2022. The estimates are retrieved from equation (C1). The average degree completion is measured as the proportion of bachelor's degree-seeking students who completed a bachelor's degree within 150 percent of the normal time (i.e., six years). The data is as of 2022 (calculated using the 2016 enrollment cohort). The increase in bachelor's degrees is calculated as the product of an increase in enrollment and average degree completion rate. For colleges of most competitive and highly competitive, the net benefit is estimated as 6.7% of the average personal and societal net benefit (i.e., $600 thousand) of college degrees. For colleges that are very competitive, the net benefit is estimated as 4.7% of the average personal and societal net benefit (i.e., $600 thousand) of college degrees.
† Mount Sinai School of Medicine does not report the degree completion data in the IPEDS.

dents from a less selective college to a more selective one instead of creating a new enrollment. Previous studies have widely established that the premium of attending a selective or elite college would exceed that of attending less selective ones (Kapur et al., 2016; Witteveen & Attewell, 2017; S. D. Zimmerman, 2019; Carnevale et al., 2022). Particularly, as demonstrated in Table E1, the majority of colleges engaged in tax avoidance behavior are categorized as most, highly, or very competitive.

I estimate the benefit of the enrollment expansion in these colleges by assuming the individual counterfactually attends a one-level lower college in Barron's Selectivity Index.[30] Specifically, for colleges categorized as most or highly competitive (tier 1 or 2), I assume that students would have

[30] The categorization is retrieved from Barron's Profiles of American Colleges. The categorization is primarily based on "college selectivity"—computed with high school performance (ranking and GPA), standardized exams, and the admission rate. It categorizes colleges into seven categories: most competitive, highly competitive, very competitive, competitive, less competitive, noncompetitive, and special (usually art or medical schools).

attended very competitive colleges (tier 3) if the colleges had not expanded their access. For colleges categorized as very competitive (tier 3), I assume that students who have attended competitive (tier 4) colleges instead. Notice that I combined the groups of most and highly competitive (tier 1 and 2) as previous studies estimated the college return based on this categorization combined the two groups and did not provide a breakdown estimation (Witteveen & Attewell, 2017).

Witteveen & Attewell (2017) estimates the earning return from most or highly competitive colleges to be 6.7% higher than degrees from very selective colleges in the short run (4 years) and 11.3% higher in the long run (10 years). Besides, the earning return from very selective colleges is 4.7% higher than attending competitive colleges in the short run and 2.1% in the long run. I treat the percentage increase in the earnings for a higher level relative to a lower level college as the premium of attending a more selective college. Then, I define the net benefit of each additional college degree granted from these colleges to be the selective premium multiplied by the estimated total personal and societal net benefits.

Step 4: Calculated the total net benefit

Combining the above statistics, I calculated the total net benefit in each college using the below formula:

$$\begin{aligned} NetBenefit_{ij} = & IncreaseEnrollment_i \times CompletionRate_i \\ & \times SelectivePremiums_j \times AvgNetBenefit \end{aligned} \tag{E1}$$

Where the net benefit of college i of selective category j is the product of the increase in degree holders ($IncreaseEnrollment_i \times CompletionRate_i$), the percentage of increase in expected earning relative to less selective colleges ($SelectivePremiums_j$), and the estimated average net personal and society benefits of a college degree ($AvgNetBenefit$). $SelectivePremiums_j$ ranges from 2.1% to 11.3% depending on the selectivity of the colleges and whether the estimation is based on the short run or long run. $AvgNetBenefit$ is obtained from previous studies, ranging from \$600 to \$1,225.

Table E1 reports the most conservative estimates based on the lowest selective premiums and total net benefits. The sum of all colleges leads to a total net benefit of $350 million. Figure E1 illustrates the ranges of estimation based on different assumptions. The estimates range from $350 million to $1,300 million.

Figure E1: Estimation of Total Net Benefit from College Enrollment Expansion