Sexual Assault on Campus: The Impact of the Women's Gymnastics Scandal on Michigan State

Kaleigh Strohl*
University of South Carolina

Bryan C. McCannon†
Illinois Wesleyan University

September 24, 2026

Abstract

In the age of social media, negative publicity can spread rapidly and damage an organization's reputation. We study how national media attention surrounding a university scandal impacts college application and attendance decisions. In December 2017, Michigan State University professor and medical doctor Larry Nassar was sentenced to 60 years in prison for assaulting over 300 women during his career as team doctor for the U.S. women's national gymnastics team and at Michigan State. The trial and conviction led to a stream of negative publicity around Michigan State for their role in the scandal. This paper studies the impact of the sexual assault scandal on university admissions. Using a synthetic control design comparing Michigan State to other similar universities, we find that Michigan State received almost 8,000 fewer applications per year as a result of the scandal, a 26% decrease. The gender-ratio of incoming students also changed, with the proportion of male students increasing by almost 2.5 percentage points. SAT Math scores also fell, with the 25th percentile and 75th percentile decreasing by 2.4% and 8.0%, respectively, suggesting a decline in student quality. These findings indicate that a high-profile institutional scandal can alter the perceptions of prospective students and impose substantial reputational costs on a university.

Keywords: college choice; higher education; institutional reputation; Larry Nassar; Michigan State University; sexual assault; synthetic control

1 Introduction

Reputation is a valuable asset for colleges and universities, impacting the number and characteristics of students they can recruit. Yet prospective students must make consequential enrollment decisions with limited information about many institutional features. While academic programs and financial aid are relatively observable, factors such as institutional governance and student safety are often more difficult to assess. Negative publicity can sharply alter perceptions of these less observable features, damaging an institution's brand when it reveals failures that conflict with students' expectations. When such information becomes public, prospective students may update their beliefs and change their college choices.

We study how adverse publicity impacts consumer choice using the public revelation of Larry Nassar's abuse and Michigan State University's handling of complaints. Nassar served as an osteopathic physician

*Ph.D. Candidate, Department of Economics, Darla Moore School of Business, 1014 Greene Street, Columbia, SC 29208, kstrohl@email.sc.edu

†Corresponding author: Dean, School of Business and Economics & Robert S. Eckley Endowed Professor of Economics, Department of Economics, 310 State Farm Hall, Bloomington, IL 61701, bryan.c.mccannon@gmail.com. We thank participants of the 2021 Southern Economic Association for helpful comments.

and professor at Michigan State and was the team physician for the U.S. women's national gymnastics team. In 2017, Nassar was sentenced to 60 years in prison for repeated sexual assault. At the same time, questions about the university's handling of prior complaints generated sustained national attention and negative publicity for the university.

Using data from the Integrated Postsecondary Education Data System (IPEDS) and a synthetic control design, we estimate the impact of the Larry Nassar scandal on applications and enrollment of MSU. The synthetic control uses a weighted combination of other universities selected to replicate Michigan State's pre-scandal outcomes and institutional characteristics. We study total applications and first-year enrollment, the gender composition of applicants and entering students, and the 25th- and 75th-percentile SAT Math scores of the entering class. These outcomes measure the degree to which the sexual assault scandal affects the decisions of potential students.

We find that the Larry Nassar scandal is associated with a decline in university demand. Michigan State received approximately 8,000 fewer applications per year than its synthetic counterpart, a decrease of about 26 percent. The composition of the entering class also changed: the proportion of male students increased by roughly 2.5 percentage points, suggesting a larger negative effect on prospective female students. In addition, SAT Math scores at the 25th and 75th percentiles declined by 2.4 and 8.0 percent, respectively, suggesting a decrease in the average quality of entering students. Overall, this evidence indicates that the high-profile nature of the scandal impacts the perceptions of prospective students; not only is there a reduction in the willingness to attend Michigan State, but the academic profile of the entering class is also weakened.

Our paper contributes to the literature on the consequences of negative publicity on university recruitment. Most relevant to this study, Johnson and McCannon [2021] find that the Penn State Jerry Sandusky child-abuse scandal worsened the academic profile of the entering class. Lawson [2021] studies the admissions consequences of sanctions against Southern Methodist University's football program, while McCannon and Wilson [2019] examine how the Duke lacrosse case affected the handling of sexual-assault cases across North Carolina. Related work links college football success to student academic performance [Lindo et al., 2012]. Many of these papers focus on misconduct or shocks tied directly to collegiate student-athletes and coaches. We extend this literature by focusing on women's safety on college campuses.

Our analysis also contributes to the literature on sexual assault in higher education, which is relatively small. Existing studies document how safety concerns reduce women's use of campus spaces [Trawalter et al., 2021] and how football-game-related partying affects reports of sexual assault [Lindo et al., 2018]. Ethnographic and interview-based research examines environments, behaviors, and beliefs associated with campus sexual assault [Armstrong et al., 2006, Schwartz et al., 2001]. We fill a gap in this literature by addressing whether a high-profile case connecting an institution to sexual assault changes prospective students' college choices. Our paper provides evidence that failures of institutional oversight can carry measurable reputational and enrollment costs.

The rest of the paper is outlined as follows: Section 2 provides background information on the Larry

Nassar scandal, Section 3 explains the synthetic control methodology, Section 4 describes data, Section 5 presents results, and Section 6 concludes.

2 Synthetic Control Method

2.1 Synthetic Control Estimation

This paper employs the approach introduced by Abadie and Gardeazabal [2003]. When attempting to study the economic impact of the rise of terrorism in the Basque Country in the mid-1970s, Abadie and Gardeazabal [2003] recognized that a simple comparison or time series estimation between the Basque Country and the rest of Spain may not capture properly the sole effect of terrorism. Potential bias may come from pre-terrorism differences in predictors of economic growth between the two groups. They approach this challenge by arguing that a weighted combination of comparison units can be better at reproducing the characteristics of the Basque Country than a single unaffected unit. This constructed "synthetic" control using other Spanish regions serves as a counterfactual for their resulting estimation. Abadie and Gardeazabal [2003] find that the rise in terrorism in the mid-1970s was associated with a 10 percentage point decline in per capita GDP in the Basque Country relative to the synthetic Basque Country. Numerous research efforts have since examined a myriad of research questions using this formalized case study approach.

Borrowing previous notation, a more formal explanation of the Synthetic Control Method can be made. Suppose there are J available control units that are unexposed to the intervention of interest (i.e., our "donor pool"). For time periods $1, ..., T_0$, this gives us $J + 1$ total units once we add in the single exposed unit (our "treated" unit). Let T_0 define the last time period before the intervention occurs, making the exposure length for the treated unit $T_0 + 1, ..., \overline{T}$.

Let $\mathbf{W} = (w_1, ..., w_j)'$ be a $J \times 1$ vector where each w_j is a nonnegative weight representing a single control unit out of the number of total unaffected J units. The vector of weights sum to one.

The Synthetic Control Method employs these weights to be able to construct a counterfactual that closely resembles the pre-intervention characteristics, or predictors, of the treated unit. We define $\mathbf{X}_1$ as a $K \times 1$ vector containing values for treatment unit variables that are potential predictors of the outcome variables of interest. Similarly, $\mathbf{X}_0$ can then be defined as a $K \times J$ matrix containing these same predictor variables, only now for the total control units J. The optimal counterfactual chosen uses the vector of weights, $\mathbf{W}^*$, which minimizes $\|\mathbf{X}_1 - \mathbf{X}_0\mathbf{W}\|$.

While we use multiple predictor variables, it cannot be assumed that these predictors are all going to influence the outcome variables by the same degree. Therefore, the $\mathbf{W}^*$ vector is going to depend on an additional weighting matrix that determines the relative importance of each of our predictor variables in reproducing the optimal synthetic. Let $\mathbf{V}$ be a $K \times K$ symmetric and positive semidefinite matrix, where the diagonal elements $(v_1, ..., v_K)$ reflect the relative importance of our predictor variables. Abadie and Gardeazabal [2003] incorporate this $\mathbf{V}$ matrix by defining $\|\mathbf{X}_1 - \mathbf{X}_0\mathbf{W}\|$ as $\sqrt{(\mathbf{X}_1 - \mathbf{X}_0\mathbf{W})'\mathbf{V}(\mathbf{X}_1 - \mathbf{X}_0\mathbf{W})}$. We can see that $\mathbf{V}$ is going to directly influence $\mathbf{W}^*$.

The goal of the synthetic's construction is to be able to estimate the outcome variable path for the treated unit if the intervention were to not occur. Let Y_{1t}^0 be this counterfactual outcome for the treated unit at time t and Y_{1t}^1 be the observed outcome. Therefore, the method attempts to reproduce Y_{1t}^0 by $Y_t^0 = \sum_{j=1}^{J} w_j^* Y_{jt}$, with j as each control unit. The estimated average treatment effect of the treated (ATT) can then be defined as $Y_{1t}^1 - Y_t^0$ for periods $t > T_0$.

The use of Synthetic Control to analyze crime is relatively new. As prominent examples, Saunders et al. [2015] considers the impact of place-based policing in an illegal drug market. Cunningham and Shah [2018] uses it to assess the decriminalization of prostitution in Rhode Island. Goin et al. [2017] considers the impact of a rainfall drought in California on crime. We use it here to assess the treatment effect of a high profile scandal on college demand.

2.2 Statistical Inference

The estimated effect cannot be fully interpreted without addressing the possibility that the synthetic control inaccurately predicted the counterfactual outcome variable path for time periods following the intervention. Abadie and Gardeazabal [2003] address this threat to the validity of the treatment effect by performing a placebo test. They create a synthetic for each control unit not exposed to the intervention, but with relatively similar characteristics to the treated unit. When comparing the outcome paths of the new synthetic and its associated control unit, a large difference between the two would indicate that the initial treatment effect estimate for the treated unit cannot be attributed to the impact of the intervention alone.

Abadie et al. [2010] build on these techniques by proposing an additional inferential method. Motivated by the idea of permutation inference, a synthetic control is created for every single unaffected unit in the sample. These synthetics create a distribution of placebo effects and it is then observed whether the magnitude of the real treatment effect can be easily reproduced. If the effect is unusually large, the probability of obtaining this estimate due to a randomly assigned intervention is $\frac{1}{J+1}$ [Abadie et al., 2010].

3 Data

All data is collected from the Integrated Postsecondary Education Data System (hereafter IPEDS).[1] We focus on the 2001 to 2019 time period as our outcome variables of interest are reported consistently since 2001. Regarding the donor pool, our primary analysis combines two lists of universities. First, we use the list of all schools which have women's gymnastics. Since the Michigan State scandal involved the doctor for the women's gymnastics team, using colleges that have women's gymnastics is one appropriate pool of schools. In addition, Michigan State is a large, public university. Our second list attempts to identify similar large institutions. We use those universities who compete in the five major athletic conferences. Combining these two lists makes up our donor pool of universities.

[1] https://nces.ed.gov/ipeds/datacenter/InstitutionByName.aspx?goToReportId=1; downloaded in August 2021.

Regarding outcome variables, IPEDS provides the 25th percentile and the 75th percentile scores for entering first-year students on both the SAT Math and the SAT Reading assessment. Coverage of the ACT scores are too limited to use in the analysis. In addition, IPEDS provides the total number of applications and the total enrollment for first-year students. Further, these two values are decomposed by gender. Hence, along with the totals, we calculate the proportion of the applications and enrollment that is male.

Missing values do occasionally arise. If a university does not provide SAT score data consistently, that school is removed from the analysis. A total of 15 universities must be removed because of missing SAT score information. On the other hand, if it is one (or just a few) year(s) with missing information, a linear interpolation is used to fill in the missing observation. For example, Louisville did not report the SAT scores of their entering first-year students in 2016. The SAT Math scores at the 25th percentile, for example, was 510 in 2015 and 530 in 2017. Thus, the linear interpolation uses a score of 520 for 2016.[2]

Regarding the predictor variables, we use a number of measurements to capture important dimensions of an institution of higher education. For the financial health, we use the price for in-state students and the price for out-of-state students. These both include not only the tuition but estimated room and board expenses. The price variables are reported for the full 2001-19 time period. In addition, we use the university's endowment. Nearly full coverage is available for the 2003-18 time period.

Regarding the student body, we use the total undergraduate enrollment and the percentage of the undergraduate student body who is male as predictors. These are, again, available for the full time period studied (2001-19). We also use the proportion of the undergraduate student body who is White (versus non-White), the percentage that reports to be of Hispanic descent, and the proportion of the student body who are foreign born (nonresident alien). These three measures are available over the 2010-19 time period.

Regarding faculty and the academic environment, we collect the faculty size (available over the 2012-19 time period), student to faculty ratio (available over the 2008-19 time period), and the undergraduate student retention rate (available over the 2004-19 time period).

As with the outcome variables, linear interpolations are used to fill in any missing values. Two additional universities are removed from the data set due to a large number of missing observations. One does not provide information on price (Air Force), while another does not report its endowment (Colorado). Thus, the donor pool used in our analysis consists of 95 universities that either have women's gymnastics and/or compete in one of the five power conference in football and basketball.[3]

Table 1 provides the descriptive statistics pooling Michigan State with the other 95 universities used in the donor pool over the years with available information.

[2]When the year 2001 is missing, the values for the year prior are used.
[3]These are the Atlantic Coast Conference (ACC), Southeastern Conference (SEC), Pac-12, Big-12, and Big Ten.

Table 1: Descriptive Statistics

	μ	σ	min.	med.	max.	coverage
Outcome Variables						
SAT Math: 25th Percentile	550.15	77.29	370	540	760	2001-19
SAT Math: 75th Percentile	658.07	71.68	470	650	8900	2001-19
SAT Reading: 25th Percentile	533.64	71.43	360	520	730	2001-19
SAT Reading: 75th Percentile	639.54	64.34	480	630	800	2001-19
Applications	19,866.1	13,646.3	670	17,525.5	113,754	2001-19
Applications: Gender Distribution	0.4553	0.0679	0.0323	0.4580	0.7316	2001-19
First-Year Enrollment	3652.8	2096.0	128	3443	15,989	2001-19
Enrollment : Gender Distribution	0.4593	0.0724	0.0226	0.4649	0.9972	2001-19
Predictors						
Division	1.156	0.488	1	1	3	2016
Power Football Conf.	0.625	0.487	0	0	1	2016
Undergraduate Student Body	18,865.1	10,566.7	481	18,432.5	76,099	2001-19
Gender Distribution	0.4710	0.0669	0.0517	0.4755	0.7215	2001-19
Price: In-State	28,861	15,343	10,251	23,837	78,654	2001-19
Price: Out-of-State	38,653	13,208	11,870	36,926	78,654	2001-19
Endowment	1,609,429,346	3,215,104,851	1,199,785	472,000,206	29,444,936,000	2003-18
Retention Rate	86.59	8.67	47	87	100	2004-19
Student-Faculty Ratio	16.19	4.47	5	17	35	2008-19
Faculty Size	1605.6	1093.6	56	1388.5	6787	2012-19
White (%)	0.6241	0.1561	0.1419	0.6594	0.8711	2010-19
Hispanic (%)	0.0933	0.0641	0.0003	0.0717	0.3450	2010-19
Foreign (%)	0.0564	0.0411	0	0.0454	0.1747	2010-19

Descriptive statistics provided for the time range considered (both pre-scandal and post-scandal years). All 96 universities, the 95 from the donor pool and Michigan State, included.

4 Results

4.1 Estimating the Synthetic

The first step is to identify the synthetic Michigan State from the donor pool. We use 2017 as the treatment year so that the matching is done for the years up to 2016. As stated, for the athletic division and the power football conference indicator variables, only one year in the pre-treatment period is used in the creation of the synthetic.

One concern that arises is that using the full coverage of the outcome variables in the identification of the synthetic can lead to the predictors being crowded out. A common practice is to selectively use the outcome variables in just a few years. We have eight outcome variables of interest and sixteen pre-scandal years. Thus, each outcome variable is used twice and only one outcome variable is used for each year.[4]

With numerous outcome variables a concern is that creating one synthetic for Michigan State will not do a good job of replicating the university in the pre-scandal years for all of the outcome variables. It is unlikely that the synthetic that best describes Michigan State's applications and enrollment will necessarily match it in SAT scores as well.

[4]We use the order of SAT Math 25th percentile, SAT Math 75th percentile, SAT Reading 25th percentile, SAT Reading 75th percentile, total applications, gender distribution of applications, total enrollment, and gender distribution of enrollment for the years 2001-08 and 2009-16.

To deal with this concern, we create eight synthetic Michigan States. The same donor pool and predictors are used for each. The eight are distinguished in that each uses every pre-scandal year for one of the outcome variables. For example, one synthetic Michigan State will be created to study applications. For it, all predictor variables will be used and each of the other seven outcome variables will be used in two years as described. It differs in that the application data is used in every pre-scandal year to determine the synthetic.

Table 2 summarizes the outcomes of the eight synthetic Michigan State's created. Table 2 lists the universities that receive a non-zero weight for at least one synthetic. It provides the number of times (out of eight) the school was selected, and the average weight used across the eight synthetics.

Table 2: Synthetic MSU

School	# times in synthetic	avg. weight
Indiana	6	0.2456
Purdue	6	0.2225
Penn State	5	0.1329
Washington	5	0.1096
Texas Women's	6	0.0540
Wisconsin	2	0.0393
Iowa	3	0.0380
Sacramento State	3	0.0241
Syracuse	1	0.0220
West Virginia	1	0.0203
Washington State	1	0.0188
Ohio State	2	0.0180
Oregon	2	0.0154
Texas A&M	2	0.0148
Long Island	1	0.0103
Kent State	1	0.0081
Michigan	2	0.0060
Yale	1	0.0003

Some universities, such as Indiana, Purdue, Penn State, and Washington, are commonly making up the synthetic and provide a large weight to its creation. Other universities are only needed once or twice, such as Yale, and contribute only a little bit to the synthetic Michigan's State creation.

For the most part, the list seems reasonable. Even though the matching did not use the identity of the athletic conference the school compete in, the synthetic Michigan State is made up primarily of geographically close universities that compete in the same conference: Indiana, Purdue, Penn State, Wisconsin, Iowa, Ohio State, and Michigan. Some schools may be a bit of a surprise. Texas Women's University is used to represent Michigan State across all SAT score outcomes and total applications, but unsurprisingly is not used to replicate the gender distribution of the first-year enrollment. In fact, five universities that are only used once arise in the synthetic which attempts to match Michigan State's gender distribution in first-year student enrollment.

Table 3 compares the mean value of the eight outcome variables for the pre-scandal years at Michigan State to the mean value of the synthetic control for that outcome variable.

ht

Table 3: Comparing MSU to Its Synthetic

	Actual MSU	Synthetic MSU
Applications	27,496.0	26,939.2
Applications: Gender Distribution	0.4669	0.5013
First-Year Enrollment	7642.0	7124.3
Enrollment: Gender Distribution	0.4591	0.4731
SAT Math: 25th Percentile	466.3	530.4
SAT Math: 75th Percentile	606.9	661.2
SAT Reading: 25th Percentile	530.3	496.0
SAT Reading: 75th Percentile	530.3	613.3

The average value for each outcome variable over the pre-scandal years is reported.

The synthetic Michigan State does a fairly good job of capturing the typical number of applications, the gender distribution in enrollment of first-year students, and the SAT Math scores in the years prior to the scandal. It will be important to recognize how well the synthetic matches the actual Michigan State in the years prior to the scandal when evaluating the divergence between the two afterwards.

4.2 Main Results

Figure 1 graphically depicts the actual Michigan State and its synthetic for the four SAT related measurements.

Regarding the SAT Reading scores of entering first-year students, Michigan State experiences a dip in the 2013-15 years that is not replicated by the synthetic. Thus, one should be skeptical that any divergence experienced after the scandal truly reflects a change in the quality of the entering student body.

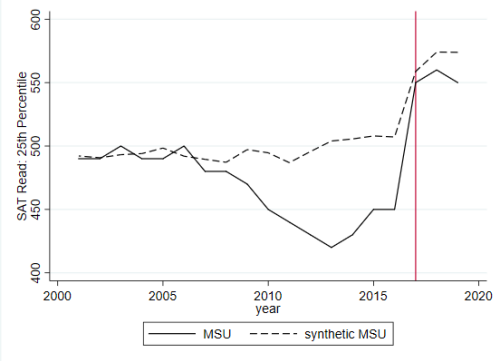
The SAT Math scores, on the other hand, do not suffer from this deficiency. The synthetic and actual Michigan State are quite close in every year prior to the scandal. For each, a noticeable divergence arises once the scandal hits. The synthetic Michigan State continues on its gradual upward trajectory, but the actual Michigan State flat-lines and, for the high scoring students, actual drops. Thus, the bottom two panels of Figure 1 strongly suggest a negative impact of the scandal on the university.

Similarly, Figure 2 considers the four outcome variable related to the quantity of students.

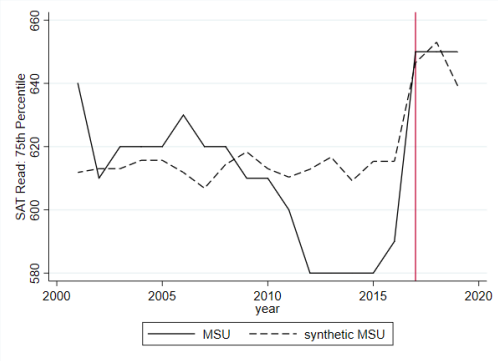
Regarding applications, the synthetic closely matches the actual and a noticeable divergence occurs. Once again, though, for the gender distribution of applications and the enrollment of first-year students the synthetic does not match the actual Michigan State very well prior to the scandal. The gender distribution in enrollment does provide an intriguing result. While the two co-move, the proportion of the entering first-year class that is male sits below the synthetic in the 2000s, but overtakes the synthetic after the scandal. This is evidence that the scandal discouraged women from enrolling.

While Figure 1 and 2 depict the time series graphically, we want to estimate an average treatment effect. To do so, we pool Michigan State with its synthetic for each outcome variable over the 2001 to 2019 time

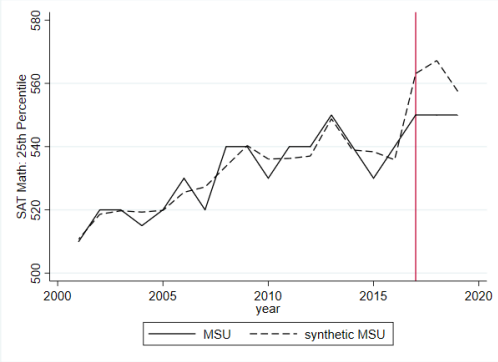
Figure 1: SAT Scores



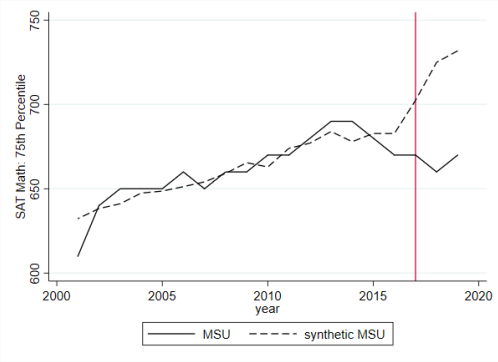
(a) SAT Reading: 25th Percentile



(b) SAT Reading: 75th Percentile

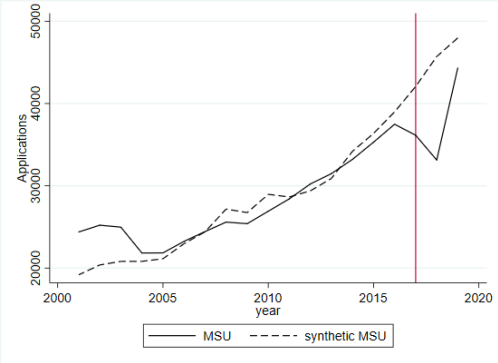


(c) SAT Math: 25th Percentile

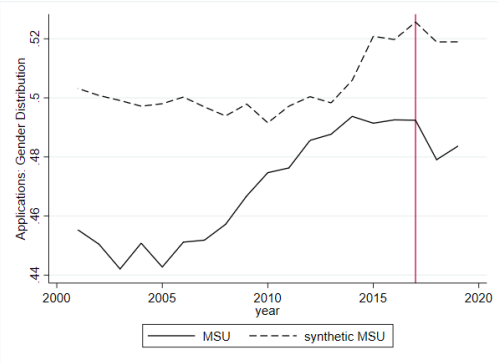


(d) SAT Math: 75th Percentile

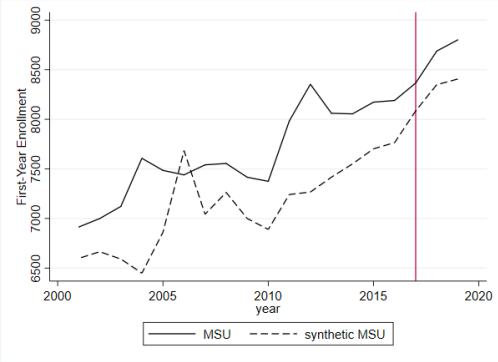
Figure 2: First-Year Students



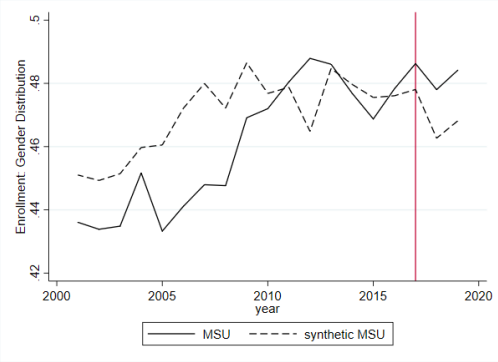
(a) Applications



(b) Applications: Gender Distribution



(c) First-Year Enrollment



(d) Enrollment: Gender Distribution

period creating a simple panel data set. With it, we estimate the following econometric model:

$$Y_{ay} = \alpha_0 + \alpha_1 Post_y + \alpha_2 MSU_a + \alpha_3 Post_y \times MSU_a + \epsilon_{ay}. \quad (1)$$

Each outcome variable, Y_{ay}, will be considered separately. The indicator variable $Post_y$ is equal to one if the observation occurs in the post-scandal years ($y \geq 2017$) and the indicator variable MSU_a is equal to one if the observation comes from the actual Michigan State, rather than its synthetic. The coefficient α_3 is the difference-in-difference coefficient of interest. It measures the average treatment effect of the scandal on the outcome variable of interest. Table 4 presents these DiD estimates.

Table 4: Difference-in-Difference Results

	Applications	Enrollment: Gender	SAT Math: 25th Percentile	SAT Math: 75th Percentile
Post x MSU	-7940.91 ***	0.024 ***	-12.51 **	-53.14 ***
	(3795.59)	(0.007)	(4.68)	(10.65)
R^2	0.536	0.182	0.496	0.457

The panel covering 2001-19 for both the actual and synthetic Michigan State is considered ($N = 38$). Only the difference-in-difference coefficient is presented. The four outcome variable shown to have a divergence in Figure 1 and 2 are presented. Robust standard errors are presented in parentheses; *** 1%, ** 5%, * 10% level of significance.

The results in Table 4 provide estimates for the scandal's effect. Michigan State experienced almost 8000 fewer applications per year than it would have had the scandal not occurred. This represents an almost 1.7 standard deviation decrease in the total number of applications received. Rather, this is a 25.6% decrease in the mean.[5] The proportion of the first-year entering class that is male increased by almost two and a half percentage points. This represents an increase of 1.3 standard deviations, or rather, a 5.1% increase in the proportion of entering students that are male. For the SAT Math scores, the distribution shifts to the left. The 25th percentile score drops by 12.5 points and the 75th percentile drops by over 53 points. These are 1.1 and 2.6 standard deviations decreases, respectively (rather, reductions in the mean SAT Math score by 2.4% and 8.0%, respectively). Thus, these are all rather substantial changes to the university's student body. We do not present the other four outcome variables as we are unable to construct a synthetic that matches the actual unit sufficiently close.

4.3 Placebo Tests

The previous section identified the synthetic Michigan State and compared it to the actual Michigan State across the eight outcome variables. The difference-in-difference results presented in Table 4 show that for four of the outcome variables the discrepancy between the two grows after after the scandal. This difference is statistically significant.

What these results do not establish is that the gap realized actually differs from what would arise by random chance, given the estimation method employed. The synthetic is constructed to match in the pre-

[5] We use the mean and standard deviation of the sixteen pre-scandal years in this calculation.

Table 5: Results from Placebo Tests

Outcome variable	Fisher exact p-value
SAT Math: 25th Percentile	0.125
SAT Math: 75th Percentile	0.042
Applications	0.073
Enrollment: Gender Distribution	0.729

For each outcome variable 95 placebo tests were conducted (one for each school in the donor pool). For each outcome variable a distinct synthetic is created using every pre-scandal year of that outcome in the determination of the synthetic. The Fisher exact p-value identifies the proportion of placebos that experience a greater divergence after 2017, relative to the RMSPE prior, than Michigan State does.

scandal years. Consequently, it is not surprising that a divergence occurs after the scandal (it is, of course, noteworthy that actual Michigan State consistently becomes worse). A placebo test can identify whether the divergence observed is unusual for the data.

In a placebo test, each university in the donor pool is considered separately. We presume that the donor pool school experiences a scandal in 2017. We identify its synthetic and measure the divergence between its actual value and its synthetic value after 2017. Specifically, we calculate the Root Mean Squared Prediction Error (hereafter RMSPE) in the years after the scandal, and divide it by the RMSPE prior to the scandal. This ratio normalizes the gap by the closeness of the match prior to treatment. A value of two, for example, is interpreted as the divergence after the scandal tends to be twice as big as it was prior to the scandal. We then iterate this process for every university in the donor pool. This creates a distribution of the RMSPE ratios that would occur by random chance. As the final step, we compare the ratio of the prediction errors for Michigan State, which did experience the scandal, to the distribution of prediction errors that arise from all of the donor pool schools. The proportion of observations with a ratio greater than that experienced by Michigan State becomes the created p-value to be used in statistical inference.

Since we allow the creation of the synthetic to vary by the outcome variable of interest, this placebo test must be re-estimated for each of the eight outcome variables. Table 5 presents the Fisher exact p-value estimated for each.

The placebo test provides supporting evidence that a nonrandom divergence has occurred. Most notably, the SAT Math score, measured at the 75th percentile of the first-year student's distribution, experiences a statistically strong effect. The exception is the gender distribution of first-year enrollment. As illustrated in Figure 2, the Synthetic Control Method creates only a poor fitting synthetic Michigan State in the pre-scandal years. Hence, the divergence observed, relative to the pre-scandal goodness of fit, is relatively weak and many donor pool universities have at least as large of a measured divergence.

5 Conclusion

We provide evidence that sexual assaults on campus affect interest in a higher education institution by evaluating the Larry Nassar scandal and its effect on Michigan State. Using the Synthetic Control Method, we document a reduction in the total number of applications, a shift in the gender distribution of first-year students, and a reduction in SAT Math scores for entering students corresponding to the scandal's timing. Therefore, the harm to the university is not only the legal expenses and harmed reputation, but the scandal affected the recruiting of new students.

While we document a reduction in SAT scores of first-year students, the data we have available to us does not differentiate scores by gender. We find some evidence that the gender distribution of first-year student enrollment shifted away from females. This affect could have been mitigated by accepting more marginal female applicants. Without being able to observe test score differences, this is left as an untested hypothesis. Further, information on the actual price students pay, after scholarships, grants, and price reductions, is not available. It would be interesting to identify what the price/revenue effects are. This would complement our analysis on the "quantity" impact we document. Nevertheless, our results provide the first measurement of adverse effects of sexual assault scandals on a higher education institution's demand.

Our results highlight that university administrators should be motivated to monitor employee behavior and respond quickly to allegations as the health of the higher education institution depends in part on the safety of women on the campus.

References

A. Abadie and J. Gardeazabal. The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1):113–132, 2003.

A. Abadie, A. Diamond, and J. Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505, 2010.

E. A. Armstrong, L. Hamilton, and B. Sweeney. Sexual assault on campus: A multilevel, integrative approach to party rape. *Social Problems*, 53(4):483–499, 2006.

S. Cunningham and M. Shah. Decriminalizing indoor prostitution: Implications for sexual violence and public health. *Review of Economic Studies*, 85(3):1683–1715, 2018. doi: doi.org/10.1093/restud/rdx065.

D. E. Goin, K. E. Rudolph, and J. Ahern. Impact of drought on crime in california: A synthetic control approach. *PLOS One*, 12(10):e0185629, 2017.

C. Johnson and B. C. McCannon. Athletics and admissions: The impact of the penn state football scandal on student quality. *Journal of Sports Economics*, 1(1):1–12, 2021.

K. Lawson. The lasting impact of ncaa sanctions: Smu and the death penalty. *Journal of Sports Economics*, 1:1–2, 2021.

J. M. Lindo, I. D. Swensen, and G. R. Waddell. Are big time sports a threat to student achievement? *American Economic Journal: Applied Economics*, 4(4):254–274, 2012.

J. M. Lindo, P. Siminski, and I. D. Swensen. College party culture and sexual assault. *American Economic Journal: Applied Economics*, 10(1):236–265, 2018.

B. C. McCannon and M. Wilson. "a million dollars in free advertising": Politics and sex offense prosecution in the wake of duke lacrosse. *Supreme Court Economic Review*, 27:167–187, 2019.

J. Saunders, R. Lundberg, A. A. Braga, G. Ridgeway, and J. Miles. A synthetic control approach to evaluating place-based crime interventions. *Journal of Quantitative Criminology*, 31(490):413–434, 2015. doi: doi.org/10.1007/s10940-014-9226-5.

M. D. Schwartz, W. S. DeKeseredy, D. Tait, and S. Alvi. Male peer support and a feminist routing activities theory: Understanding sexual assault on the college campus. *Justice Quarterly*, 18(3):623–649, 2001. doi: 10.1080/07418820100095041.

S. Trawalter, J. Doleac, L. Palmer, and A. Carter-Sowell. Women's safety concerns and academia: How safety concerns can create opportunity gaps. *Social Psychology and Personality Science*, 1:1–2, 2021.

6 Appendix

Table 6: Schools Dropped From the Donor Pool

Missing SAT Information	
Alaska - Anchorage	Wake Forest
Gustavus Adolphus	Western Michigan
Illinois State	Winona State
Ithaca	Wisconsin - Eau Claire
Kansas	Wisconsin - Lacross
Kansas State	Wisconsin - Oshkosh
Oklahoma	Wisconsin - Whitewater
Southeast Missouri State	
Missing Price	**Missing Endowment**
Air Force	Colorado

Table 7: Schools Making Up the Synthetic Michigan State

	Total Applications	Applications: Gender	First-Year Enrollment	Enrollment: Gender	SAT Math 25th Percentile	SAT Math 75th Percentile	SAT Reading 25th Percentile	SAT Reading 75th Percentile
Indiana	0.288	0.332	0.553	0.227	0.283	0.282	∅	∅
Purdue	0.344	0.292	0.236	∅	0.254	∅	0.537	0.117
Penn State	0.043	0.014	0.010	∅	0.247	∅	∅	0.749
Washington	0.026	0.058	0.033	∅	∅	0.602	0.158	∅
Texas Women's	0.066	0.076	∅	∅	0.055	0.018	0.084	0.133
Wisconsin	0.201	0.113	∅	∅	∅	∅	∅	∅
Iowa	∅	0.049	∅	0.137	0.118	∅	∅	∅
Sacramento State	∅	∅	∅	0.001	0.042	∅	0.150	∅
Syracuse	∅	∅	∅	0.176	∅	∅	∅	∅
West Virginia	∅	∅	∅	0.162	∅	∅	∅	∅
Washington State	∅	∅	∅	0.150	∅	∅	∅	∅
Ohio State	∅	0.066	0.078	∅	∅	∅	∅	∅
Oregon	∅	∅	∅	∅	∅	0.099	0.024	∅
Texas A&M	∅	∅	0.074	∅	∅	∅	0.044	∅
Long Island	∅	∅	∅	0.082	∅	∅	∅	∅
Kent State	∅	∅	∅	0.065	∅	∅	∅	∅
Michigan	0.032	∅	0.016	∅	∅	∅	∅	∅
Yale	∅	∅	∅	∅	∅	∅	0.002	∅

Each column provides the weights used to create the synthetic Michigan State. Results differ as each synthetic's creation includes each year of the pre-treatment period for the outcome variable selected. Each other outcome variable is only used in two of the sixteen pre-treatment years.

Table 8: Prediction Errors

	Total Applications	Applications: Gender	First-Year Enrollment	Enrollment: Gender	SAT Math 25th Percentile	SAT Math 75th Percentile	SAT Reading 25th Percentile	SAT Reading 75th Percentile
Applications	**2282.68**	2751.06	3355.63	9311.94	3015.92	4817.72	4593.82	6696.97
	8299.00	6533.68	9230.94	11068.05	9364.43	6221.02	5785.00	22741.79
	3.636	2.375	2.751	1.189	3.105	1.291	1.259	3.396
Apps: Gender	0.046	**0.038**	0.032	0.021	0.043	0.022	0.076	0.057
	0.047	**0.036**	0.028	0.035	0.029	0.028	0.072	0.016
	1.016	**0.959**	0.875	1.662	0.671	1.248	0.941	0.286
Enrollment	1361.38	1455.54	**603.55**	3207.34	1471.92	1844.64	2083.98	1587.67
	1139.43	1355.43	**341.49**	3559.16	1176.95	1691.36	1817.67	2755.51
	0.837	0.931	**0.566**	1.110	0.800	0.917	0.872	1.736
Enroll: Gender	0.054	0.049	0.047	**0.018**	0.053	0.019	0.083	0.060
	0.034	0.027	0.036	**0.014**	0.034	0.020	0.042	0.034
	0.619	0.551	0.768	**0.768**	0.632	1.055	0.502	0.573
SAT Math: 25	17.78	10.755	11.10	23.37	**4.20**	14.92	7.22	12.48
	41.82	35.16	34.81	3.21	**13.22**	40.88	22.35	16.99
	2.352	3.270	3.136	0.137	**3.15**	2.740	3.094	1.361
SAT Math: 75	11.00	10.61	11.06	42.75	16.14	**8.43**	15.98	19.04
	47.37	40.94	41.85	15.09	9.30	**55.07**	36.89	15.88
	4.306	3.859	3.782	0.353	1.196	**6.531**	2.309	0.834
SAT Read: 25	56.09	50.71	57.17	31.95	45.44	56.67	**42.55**	51.82
	31.94	25.37	29.03	4.32	10.86	31.57	**16.82**	11.98
	0.569	0.500	0.508	0.135	0.239	0.557	**0.395**	0.231
SAT Read: 75	29.93	28.14	32.33	18.04	23.20	36.11	24.80	**20.58**
	25.10	21.83	26.48	4.35	8.76	32.61	18.60	**6.78**
	0.839	0.776	0.819	0.241	0.378	0.903	0.750	**0.329**

Each column is the results from the synthetic control that target that outcome variable in the creation of the synthetic. Each collection of three row is the outcome variable of interest. Within each cell, the top number is the Root Mean Squared Prediction Error for the pre-scandal periods, the middle number is the Root Mean Squared Prediction Error for the post-scandal periods, and the bottom number is the ratio of the two. The bolded values depict the values of the synthetic used in the paper.