Project Studies

University Ranking Effects: A Descriptive Analysis

Chris Vennemann
c_venn01@uni-muenster.de
Matr. No.:429793

Institute for Public and Regional Economics
Summer Term 2021

Contents

List of Tables

LIST OF FIGURES

1 INTRODUCTION

This paper analyses longitudinal data on university rankings and a range of other university characteristics like size, selectivity, and costliness. Yearly administrative data from the US College Scorecard, provided by the US Department of Education, and historical data on US-college rankings allow for an analysis of time-dynamics from 1996 to 2018.

The objective of this paper is to give an overview of important variables and to explore the characteristics inherent to the data. Specifically, I analyze how the data varies across universities and over time. To do so, I group similar universities into clusters generated by an unsupervised clustering algorithm. This clustering approach primarily aims to simplify the analysis of heterogeneity across universities as comparing single universities can become troublesome—even with the moderately low number of 50 universities in the sample. The merit of clustering lies in the nature of clustering algorithms because clusters are generated such that they have low within but high between variation. Hence, they may offer benefits over manually formed groups, especially since it is hard to know a priori which universities are similar and which are not. Throughout this paper, the variables are analyzed with respect to a proceeding empirical analysis. To this end, I give a more succinct overview of variables that I plan to use as control variables and a rather comprehensive breakdown of dependent variables and the ranking variable. A separate chapter will be devoted to the ranking variable as it stands out by having partially very little variance throughout the years.[1] This strong persistence of the rankings needs to be accounted for in a proceeding empirical analysis in order to identify causal effects. This likely concerns, but is not limited to, the decision whether to use lagged rankings, same-period rankings or both in an econometric model.

The remainder of this paper is structured as follows. Section 2 forms the main part of this study. Initially, the data generation is briefly discussed. In the following subsections, I give a detailed overview of variable definitions and their measurements and make a selection of important variables which are more thoroughly studied in the course of this section. The important variables are analyzed both in terms of hetero-

[1] Most of the universities in the sample have strongly persistent rank placements throughout the years, while others have caught up substantially in the ranks.

geneity across universities and over time. The last part of Section 2 deals with the time-persistency of the rankings. Finally, Section 3 concludes the most important insights of this paper and gives a brief outlook on potential threats of endogeneity in an econometric analysis.

2 Data & Methodology

The data was obtained from the College Scorecard, an online tool designed by the US Department of Education to assist students in making informed college choices. The College Scorecard allows anyone to compare US universities on a wide range of characteristics. These include, among others, graduation rates, expected salaries, or average costs of attendance. The data from the College Scorecard itself is a compilation of data from other administrative data sources.[2] Each of the yearly datasets contains approximately 2000 variables on around 6000 units of observations. I made a selection of around 40 variables that cover information on size, programs, costs, and student body at the university level. These are either provided by the Integrated Postsecondary Education Data System (IPEDS) or National Student Loan Data System (NSLDS). Table A1 gives an overview of descriptive statistics of these variables.

The ranking variable covers placements of the top 50 universities from the US News & World Report US Best College Rankings. This ranking was first published in 1983 and has appeared yearly since. Besides other college rankings, it is one of the most popular and anticipated rankings among students and scholars in the US. US News Rankings are divided into separate rankings for National Universities, National Liberal Arts Colleges, Regional Universities, and Regional Colleges. This study focuses on National Universities. US News describes them as offering a wide range of undergraduate, master, and doctoral programs as well as pioneering in research. On the other hand, National Liberal Arts Colleges are primarily undergraduate-focused and offer predominantly liberal arts programs. Regional Universities and Colleges are small undergraduate institutions and offer fewer master and doctoral programs.

I obtained the rankings data from Reiter 2021 who has compiled the data from historical National Universities rankings in a single dataset covering a period from as far

[2] E.g., the Integrated Postsecondary Education Data System (IPEDS), US Department of Treasury, or National Student Loan Data System (NSLDS)

as 1984 until today. This dataset contains information on ranking placements of every university ranked in the top 50 at least once from 1984 to 2021. I drop universities with more than 15 missing rank entries in the 23-year period (1996-2018). These are typically the "marginal" universities that may have placed around 50th in one year but have not reached the top 50 in other years. After all, the final panel dataset contains 50 universities over the period from 1996 to 2018.

2.1 VARIABLES

I make a selection of dependent variables covering information on universities selectivity, size, and costliness. These are a university's admission rate, average SAT of its student body, number of undergraduate and graduate students, as well as in-state and out-of-state tuition fees.

Table A1 shows summary statistics of all variables pooled over units of observation and time. C_V denotes the coefficient of variation, which is simply the quotient of the standard deviation divided by the mean. It is moderately high for the admission rate, the number of undergraduates and graduate students, and in-state tuition fees. Although this sample is already limited to National Universities, the number of undergraduate and graduate students still varies substantially between the different universities. Contrarily, the low variation of the SAT average is somewhat expected due to the nature of its measurement because the interval of possible realizations already greatly limits a possible source of variation. The later sections expand on how the dependent variables vary across units and over time. This is insofar necessary as the pooled statistics of Table A1 do not allow for qualitative assessments based on which dimension causes the most variation of the variables. Particularly Section 2.3 will show that some of the dependent variables exhibit considerable movements over time.

The control variables are divided into three subcategories. Panel A describes the student body composition and the percentages of degrees awarded in a particular field of study, denoted by the "PP" (program percentage) prefix. Overall, these variables show high variation. In Section 2.2, where clusters are introduced to study heterogeneity across universities, the percentages of degrees awarded in Computer & Information Science and Engineering play a decisive role in how the clusters are formed. The variables with the "UGDS" prefix indicate the shares of the respective ethnicity in the undergraduate student body. Unsurprisingly, the majority of undergraduates at most universities are white. The second most represented ethnic group is Asian, closely followed by Hispanic and Black. Unfortunately, these variables are only available since 2008 due to changes in ethnicity definitions—hence their low number of observations in the sample. The last four variables in panel A describe the parent's education. On average, around one in five students is a first-generation student. Two percent have parents with only middle school education, 18% with high school education, and 80% have parents with post-secondary education. Because the possible collinearity of the first three variables may pose a threat in a regression framework, dropping the first-generation variable may be advisable when controlling for school heterogeneity.

I group variables, which characterize universities on a more general level, into panel B. Control categorizes schools into public (1) and private (2) institutions. Of the 50 universities in the sample, 17 are public, and 33 are private. Since this characteristic does not change over time, it should not be included as a control variable—given that university fixed effects account for unit-specific, time-invariant factors. Full-time faculty indicates the share of full-time positions at a given university. The retention rate measures the proportion of undergraduate students who are still enrolled at the university one year after starting at the university. Completion rate 100 (150) specifies the share of undergraduate students who completed a degree within 100% (150%) of its expected time (typically four years). On a related note, the last three variables in this panel display the shares of students who completed a degree within two, three, or four years after starting at the institution. Given that the sample consists entirely of four-year institutions (i.e., those which predominantly offer four-year programs), the close means of the 100%-completion rate (0.72) and the share of students, who completed a degree within four years at the university (0.69), is not surprising.

Finally, panel C comprises information on costs, finances, and aid. I have data on the undergraduate shares that received a federal loan or pell grant in a current academic year and data on the student share that have ever received pell grants while in school. The average annual costs include tuition and fees, instructional supplies (e.g., books), and living expenses for undergraduate students who receive Title IV aid.[3] The

[3]Title IV aid covers all sorts of federal grants and loans to assist student's access to post-secondary education, e.g., Pell Grants or federal subsidized and unsubsidized direct loans.

net tuition revenue is calculated as the revenue from tuition minus discounts and allowances per student (undergraduate and graduate) at a given university. Analogously, instructional expenditure covers expenses per student made by the university.

2.2 HETEROGENEITY ACROSS UNIVERSITIES

In this section, I take a closer look at how the dependent variables vary across universities. Therefore, I group the universities into clusters using an unsupervised clustering algorithm. The findings of this section may give valuable ideas on how the rankings affect the dependent variables. For instance, the potential heterogeneity of ranking effects can be estimated by modeling interactions between cluster affiliation dummies and the ranking variable.

I use the popular K-Means algorithm.[4] Its objective is to divide the data points (the universities) into K clusters such that the total within-cluster variation of the K clusters is minimized. Here, I follow the conventional methodology and use the squared euclidean distance between the universities' standardized variable realizations to quantify variation. Broadly speaking, the algorithm groups universities into clusters based on how similar the values of their variables are. Therein lies also the advantage over manually formed groups: grouping manually may be feasible over one, two, or perhaps three variables, but the inherent heterogeneity over all variables can only accurately be reflected by a clustering approach. The resulting disadvantage, however, may be inferior interpretability. Inferences of ranking effect heterogeneity are arguably easier to make with one-dimensional groups such as by control (private vs. public), size (many students vs. few students), or rankings themselves (top 25 vs. top 50). Such methods employ, e.g., Meredith 2004 and Bowman and Bastedo 2009.

Table A2 shows how the universities are grouped together. Table A3 summarises additional information on the control variables. The overall pattern of the clusters is discernable. Cluster 1 consists of five universities which are characterized by a set of defining features. These are primarily the technical-mathematical universities that predominantly award degrees in Computer & Information Sciences, Engineering, and Maths & Statistics. On average, they award twice to four times as many degrees in these fields than do universities from the other three clusters. Furthermore, while

[4] See e.g. James et al. 2013 pp.386-390

the other clusters have a slight majority of female students, the universities from this cluster have an average female share of only 35%.

The elite universities form the second cluster. Of the 15 universities in this cluster, seven are Ivy-League institutions. Additionally, all 15 are private universities, have significantly higher instructional expenditures per student, and the most successful students as indicated by high retention rates, completion rates, and SAT averages. Naturally, the leading role is also reflected in the rankings—these universities stand out from the rest with a formidable average rank of 9 over all 23 years.[5]

I dub the third group the "second-best" universities because, over most variables, they do not have as good realizations as the elite universities but still clearly distinguish themselves from the rest. This group contains 17 universities—14 of them are private. The average rank of 28 and the dependent variable means further support the idea of second-best universities.

The universities in the fourth and last cluster are characterized by an overall high number of undergraduate students. All 13 universities are public, have on average students of lesser ability, and therefore also low completion rates. I named this cluster the mass universities due to their defining, high number of undergraduate students.

Table 1 compares the clusters in more detail for the dependent variables. The cluster pattern mentioned above is again clearly recognizable. Starting with selectivity, the elite universities have the lowest admission rate, followed by the second-best and the mass universities. The technical universities are similar to the second-bests as their respective means only differ by one percentage point. This is also supported by the p-values of pairwise t-tests. Here, the p-values in a column refer to a t-test with the column to its right, i.e., the first p-value in the first column is obtained from comparing the technical universities to the elite universities, the second is obtained from comparing the technical to the second-best universities, and so on.[6] For the SAT average, the relationship is somewhat mirrored: the elite universities have the highest performing students, those of the second best universities perform slightly worse, and the mass universities have the worst-performing students. Only now, the technical cluster is

[5] The mean value of nine of the 15 universities can be explained by the fact that ties are allowed in the US News Ranking.

[6] Table 1 reports p-values for pairwise independent sample t-tests as the sample size of the clusters is not equal. The p-values are adjusted for multiple comparisons following Benjamini and Hochberg 1995.

Table 1
Descriptive Statistics of Dependent Variables by Cluster

	Technical	Elite	Second-best	Mass
	(N=5)	(N=15)	(N=17)	(N=13)
Admission Rate				
Mean (SD)	0.32 (0.21)	0.16 (0.09)	0.33 (0.12)	0.47 (0.16)
Min., Max.	[0.06, 0.80]	[0.04, 0.46]	[0.11, 0.78]	[0.14, 0.79]
P-Values	<0.01,0.65,<0.01	<0.01,<0.01	<0.01	
SAT Average				
Mean (SD)	1420 (83.5)	1450 (42.1)	1360 (45.9)	1250 (77.1)
Min., Max.	[1290, 1570]	[1310, 1540]	[1270, 1500]	[1070, 1450]
P-Values	<0.01,<0.01,<0.01	<0.01,<0.01	<0.01	
Undergraduates				
Mean (SD)	6117 (4325)	7565 (2941)	10473 (7130)	29333 (7839)
Min., Max.	[879, 17223]	[2737, 14864]	[3036, 30715]	[13716, 43709]
P-Values	<0.01,<0.01,<0.01	<0.01,<0.01	<0.01	
Graduates				
Mean (SD)	4481 (3012)	9314 (6066)	8095 (6629)	10128 (4384)
Min., Max.	[1039, 17040]	[1388, 26463]	[1240, 34852]	[2332, 21445]
P-Values	<0.01,<0.01,<0.01	0.011,0.049	<0.01	
Tuitionfee in-state				
Mean (SD)	32495 (14820)	38347 (9211)	33662 (13544)	9154 (3895)
Min., Max.	[3454, 55465]	[16444, 59430]	[4160, 56382]	[1941, 18454]
P-Values	<0.01,0.498,<0.01	<0.01,<0.01	<0.01	
Tuitionfee out-of-state				
Mean (SD)	35798 (10391)	38347 (9211)	37466 (9583)	27007 (8596)
Min., Max.	[12350, 55465]	[16444, 59430]	[17017, 56382]	[7812, 43394]
P-Values	0.056,0.204,<0.01	0.25,<0.01	<0.01	

comparable to the elite cluster in terms of the average SAT as their mean comes close to the one of the elite universities.

Regarding the number of students, the more selective the universities in a cluster, the fewer is the average number of undergraduate students. Particularly the mass universities stand out, having almost three times as many undergraduate students as the rest. The technical institutions have, on average, the fewest undergraduates, which is not surprising given their specialization in programs. The same also applies to the number of graduate students, although the other three clusters are not as clearly

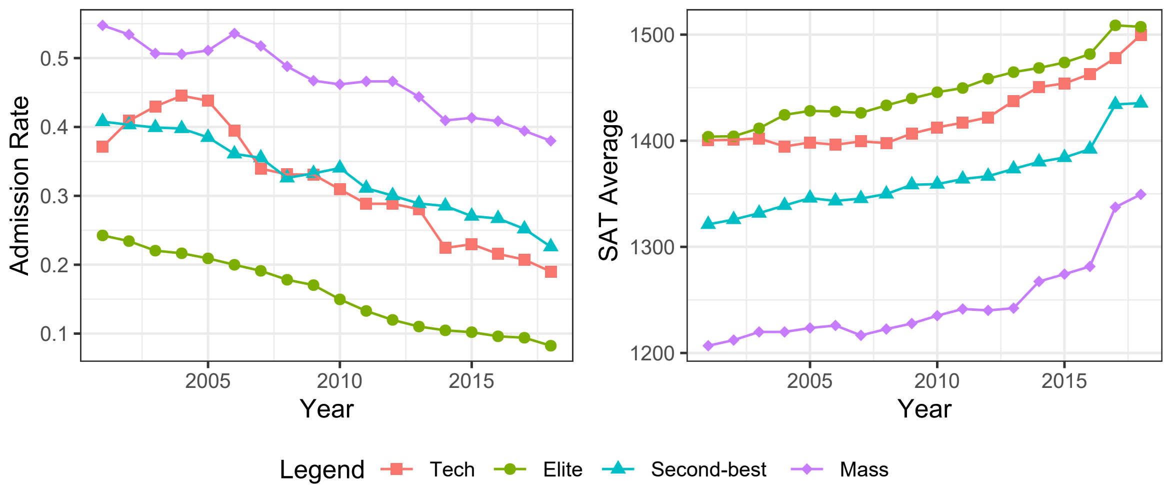

Figure 1: Time Movements of Dependent Variables by Cluster

separated from each other as with undergraduates.

Lastly, tuition fees are subject to pronounced heterogeneity across the clusters. Focussing on in-state tuition, the means increase with higher selectivity. Furthermore, the difference between the elite and second-best universities is much lower than the difference between the second-best and the mass universities. These differences are probably driven by the distinction between private and public universities since the mass universities are exclusively publically controlled, whereas universities in the other three clusters are predominantly privately controlled. The in-state tuition fee means grouped only by the control of universities is $9,460 for public and $38,426 for private institutions (not documented).

2.3 Heterogeneity over Time

This section describes the time trajectories of the dependent variables. Throughout, I keep the distinction between university clusters to show whether they move differently over time. Figure 1 displays the admission rate and the SAT average from 2001 to 2018.

The admission rate variable shows a cluster-wide downwards trend. The insights from the previous section are easily recognizable as indicated by the level of the curves: the elite schools steadily decrease on the lowest level, while the mass universities do so on the highest level. In between are the second-best and the technical universi-

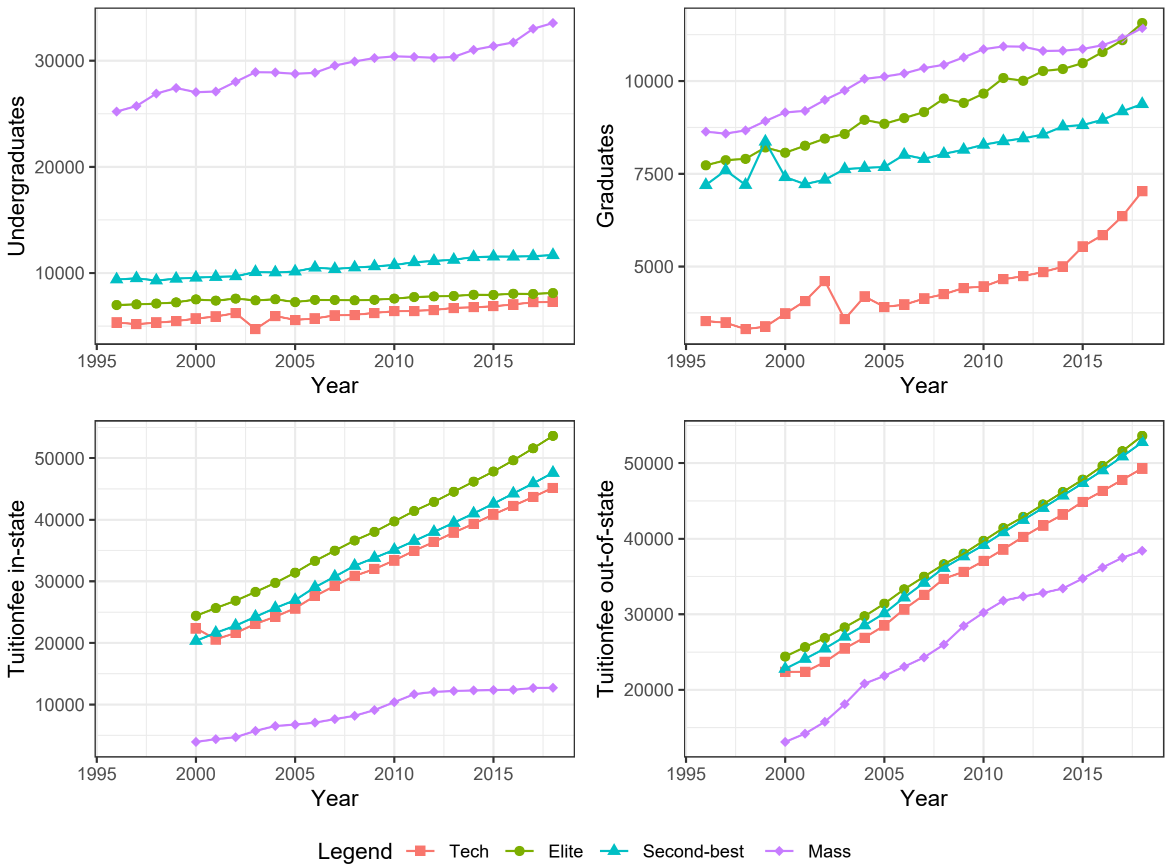

Figure 2: Time Movements of Dependent Variables by Cluster (2)

ties, which also move downwards but cross a few times. Whether this downward trend reflects that universities become more selective over the years is open for debate. Arguably, the curves may move downwards because applications increased in later years, and therefore the share of accepted students decreased.

The SAT average of the student body also depicts clearly separated curves. The mass universities fluctuate on the lowest level, followed by the second-best universities. The elite universities are on top, although the technical universities have caught up to their level in terms of student SAT achievements.

The remaining four variables are depicted in Figure 2. Unsurprisingly, the mass universities stand out by having substantially more undergraduate students than the rest. Furthermore, they have a stronger positive time trend, which is manifested in the difference in undergraduate students from 1996 to 2018. The mass universities have gained, on average, around 8,000 undergraduates over this period, while the other three groups have gained a little short of 1,800 undergraduate students on average (not documented). The number of graduate students progresses somewhat differently. Here, the cluster-wide trend is again positive and appears to be stronger than for the undergraduate variable. Technical universities are on a lower overall level, probably due to their specialized programs. Still, their curve and the curve of the elite universities rise faster than the other two clusters in more recent years. In fact, the elite universities have surpassed the level of the mass universities as of 2018 in terms of graduate students.

Regarding tuition fees, the elite, technical, and second-best universities are very similar. All three show a steady increase throughout the years, both in terms of in-state and out-of-state tuition fees. Only the group of mass universities stands out again. Their in-state tuition fee curve is considerably flatter and on a lower level than the curve of the other clusters. The main driver of variation is likely the control of the institutions, considering that public universities are more open to anyone, whereas private universities probably serve a wealthier clientele. The graph depicting the out-of-state tuition fee looks almost identical, except that the curve of the mass universities shows a slope more similar to the other clusters. Here, the graph casts doubt on whether the tuition fee variable of the elite, second-best, and technical universities reacts at all to changes in other variables, as it looks more like the tuition fee increases annually by a fixed value, irrespective of other circumstances.

2.4 Ranking Placements

In this final section, I describe the ranking placements in more detail. The first part deals with the variability of the rankings, which is small overall—although a few universities stand out. In the second part, I illustrate correlations between the rank placements and (one of) the dependent variables.

Plot (a) of Figure 3 depicts the cumulative distribution function of average rank changes per year for the 50 universities. 26% (13) of the universities have an average rank change per year of 1 or less. These correspond to the points below the 1-line in plot (b). Most of these are from the elite or second-best cluster. 82% (41) universities have gained or lost two or fewer ranks per year. As indicated by the flattening CDF curve, a small number of universities have much higher rank variability throughout the 23

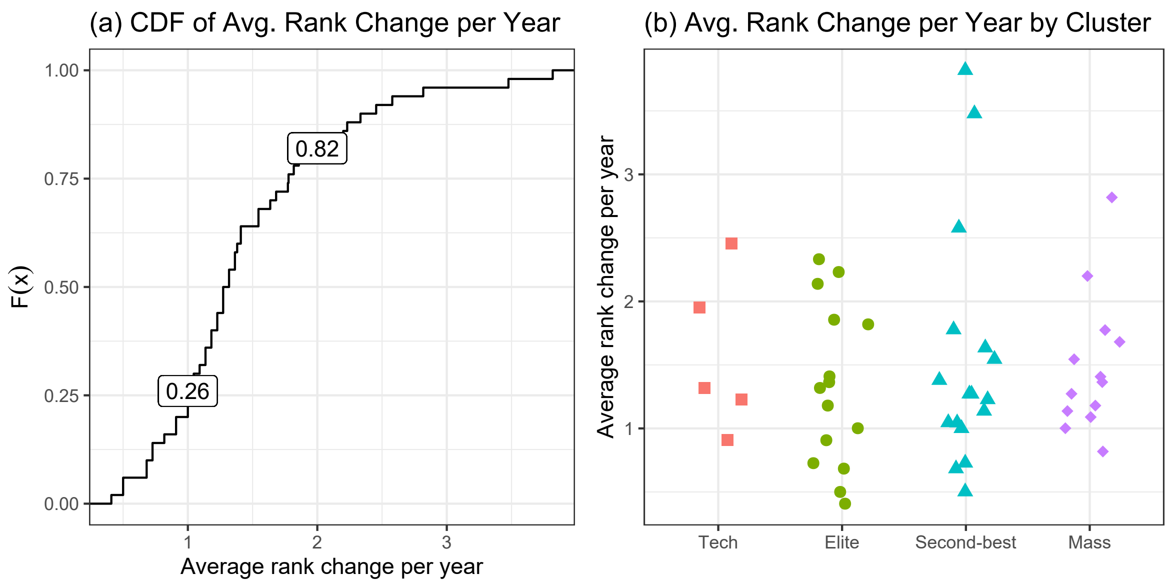

Figure 3: Average Rank Changes per Year

years, which is visible in plot (b). Two universities emerge whose ranking trajectory is discussed in more detail down below. They correspond to the two teal triangles on top of plot (b).

Figure 4 plot (a) shows the average ranking placements of the four clusters over the years. The observations from above are also evident here as the elite universities exhibit little to no variability in rankings. Over the 23 years, they have always placed 10th or below, while the other three clusters have more fluctuating curves. Especially the technical universities showed a steep drop in 2000—possibly impacted by a single university as this cluster consists of only five universities. The bottom line of this first graph is that ranking variability appears to be more pronounced at the bottom of the rankings.

Plot (b) supports this assumption. It depicts the universities with the two highest and two lowest average ranking changes per year. Here, the two highest universities are the two observations from the mass cluster mentioned earlier (University of Illinois at Urbana-Champaign and University of California-Irvine). They evidently fluctuate around the lower end of the ranking. On the other hand, the two universities with the lowest ranking variation are from the elite cluster (Harvard University and Princeton University). Since 2001, they have been sharing first and second place in the US News Best College ranking. This may indicate an incumbent effect: elite universities profit

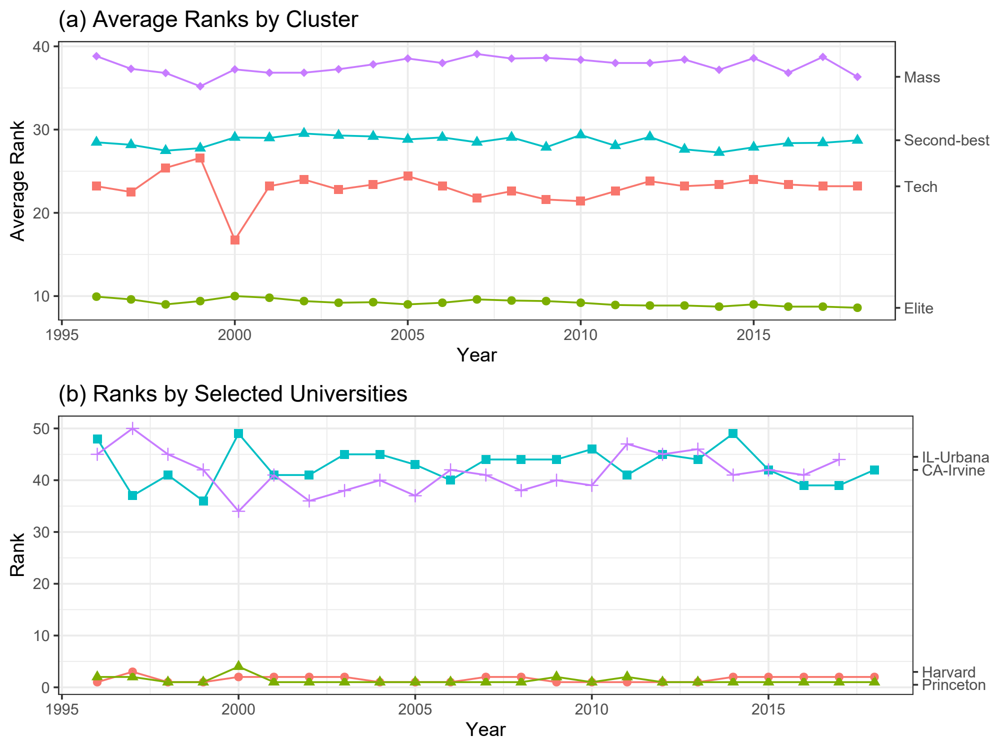

Figure 4: Ranking Placements over Time

from good ranking placements, which leads more and better students to apply to the university in following years. The universities can then choose among better students, who enhance the universities' characteristics. These characteristics (e.g., retention rates, completion rates, faculty spending per student) in turn are the basis for the universities' ranking placements in following years.

In fact, US News commonly readjusts the weights of the characteristics that play a role in the calculation of the rankings every year. As an arbitrary example, financial resources per student spent by the university contributed 10% to how universities are ranked in the 2021 edition of the National University Ranking. However, it need not have the same weight in previous years, and the same applies to any of the characteristics used. So, unless a university suddenly became so much more worthwhile than others in just one year, sudden jumps in the rankings ladder (e.g., shown by large deviations from the 45° line in the following figure) may simply reflect changes in these factor weights for the most part.

In support of the time persistency of ranking placements, Figure 5 shows how ranks

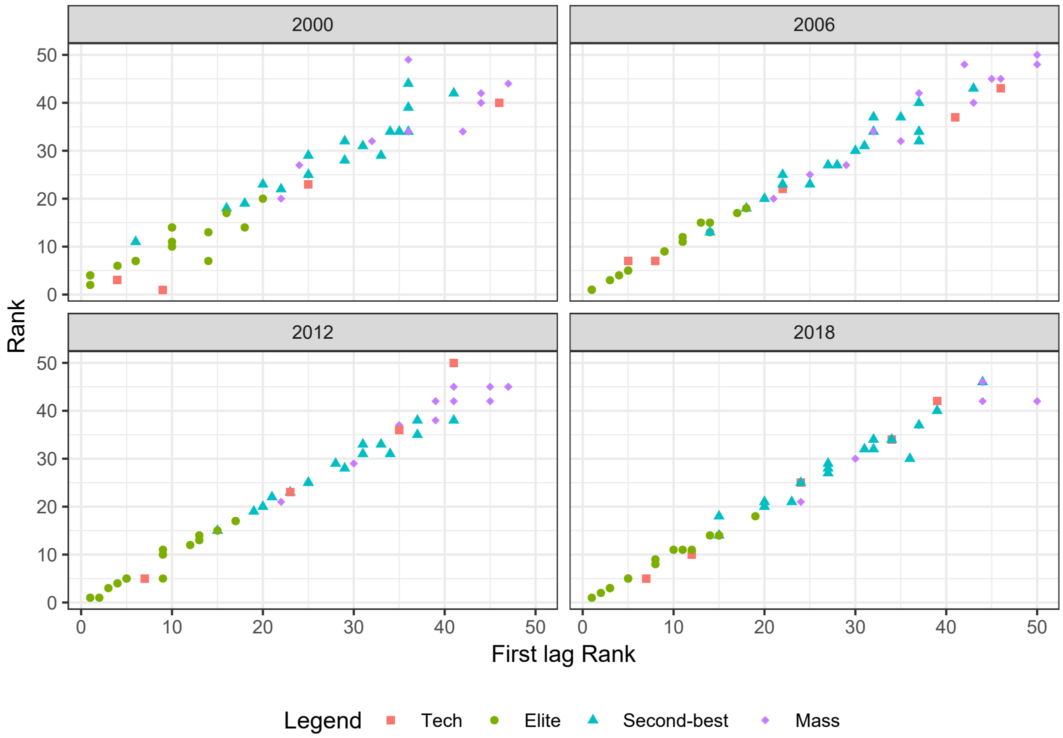

Figure 5: Correlations of Ranking Placements with First Lags

correlate with ranks of previous years. Here, each of the four graphs illustrates the relationship of the ranks with their lagged values in a specific year, i.e., the top left graph compares ranks from 2000 with ranks from 1999, the top right compares ranks from 2006 with ranks from 2005, and so forth. The correlations remain very strong over the years. Only in 2000, the observations scatter more loosely around the 45° line and therefore likely indicate considerable changes in the factor weights. There is also an observable difference at the position of the ranks: towards the right, i.e., at worse ranks, the correlation is overall weaker than at the good ranks, sustaining the observation of low variability among the best rank placements. Additionally, Figure B1 compares the relationship of current-year ranks with ranks from four years ago. It shows that the correlations of the ranks between four years are also strong and remain strong over the years.

Figure 6 shows, as an example, the relationship between the SAT average and the ranking placements. With a correlation coefficient of $r = -0.81$, the correlation is strong and negative, suggesting that higher-ranked universities are associated with

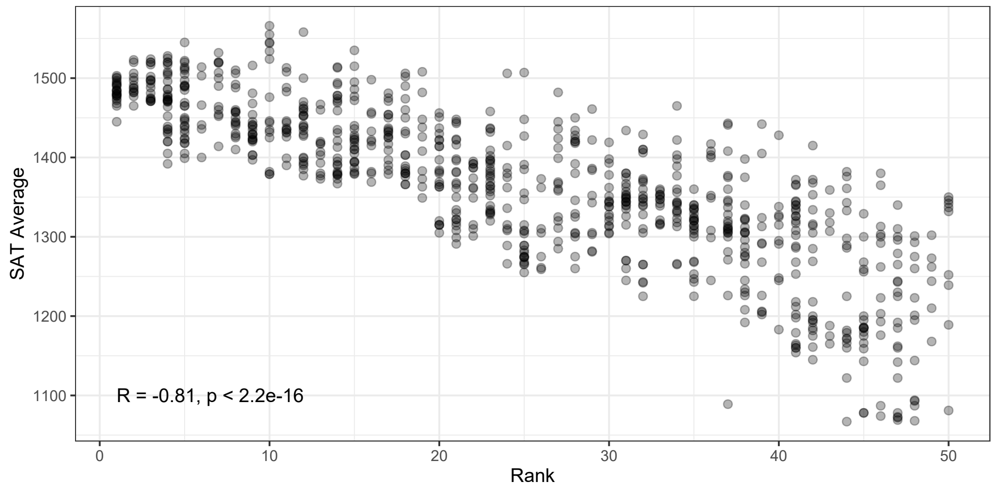

Figure 6: Correlation between SAT Average and Ranks

higher student SAT scores, on average. The figure includes observations from 2001 to 2018 because the data is not available before that period. Figure B2 additionally decomposes the observations into clusters and different years, indicating that the relationship remains similar across universities and over the years. Only in 2018, the regression slope for the elite universities is somewhat flatter than in the others graphs.

The depicted relationship leads to an important question: to what extent do the ranking placements *cause* differences in the dependent variables, e.g., the average SAT of a university? With the formulation of the research question, some concerns about the identification strategy arise. First, do other factors confound the relationship between the dependent variables and the ranks? The results from this study indicate that this is almost certainly the case since universities with good students tend to have better retention and completion rates. Because the ranking placements are in parts determined by these factors, omitting them would lead to bias in a regression framework. Secondly, the variables might be subject to simultaneous causality if the dependent variable directly determines the placements. It seems reasonable to assume that this may only be prevalent in same-year contexts as the rankings of a given year are calculated based on variable realizations from that respective year. Also, this concern is more strongly warranted for the admission rate and the SAT average because they are known to have co-determined the university rankings at some point. Thirdly, the timing of the effect is unclear. It seems plausible that, if a causal effect is present, the rank from the year before induces changes in the dependent variable—although ranks from far back years may still partially affect present-year dependent variable outcomes, too.

3 CONCLUSION

The goal of this analysis was to give a thorough overview of the data. I selected six dependent variables and the ranking variable, which were analyzed in terms of variation between universities and over time.

The clusters generated by the K-Means algorithm resulted in clearly separated groups of universities. The group of technical universities is defined by high shares of awarded degrees in Computer science, Maths, Statistics, and Engineering. In other characteristics, the technical universities are often on par either with the elite or the second-best universities. The elite cluster contains prestigious private institutions characterized by low admission rates, high achieving students, and high costs of attendance. The second-best universities build the middle of the pack. Over most variables, they score significantly below the elite universities but still higher than the last group. The fourth group compromises exclusively public universities and is therefore primarily defined by a high number of (under)graduate students and low cost.

All six dependent variables show noticeable time trends with overall similar slopes. While the shares of accepted students decreased considerably since the new millennium, the student quality, the number of students, and tuition fees all increased over the 23 years—showing only minor exceptions in trend but some stark differences in levels.

The ranking placements are persistent and highly correlated over time, few exceptions notwithstanding. Furthermore, the figures showed that persistence is more pronounced at better ranking placements.

The results also showed that other variables probably confound the relationship between the ranks and the dependent variables. This confounding can likely be resolved by appropriate control variables, provided they are included in the sample. Two other concerns, reverse causality and the correct timing of the effects, may not be so easily resolved. At the very least, they deserve further attention with respect to the appropriate identification strategy of an econometric analysis.

References

Benjamini, Yoav and Yosef Hochberg (1995). “Controlling the false discovery rate: a practical and powerful approach to multiple testing”. In: *Journal of the Royal statistical society: series B (Methodological)* 57.1, pp. 289–300.

Bowman, Nicholas A and Michael N Bastedo (2009). “Getting on the front page: Organizational reputation, status signals, and the impact of US News and World Report on student decisions”. In: *Research in Higher Education* 50.5, pp. 415–436.

James, Gareth et al. (2013). *An introduction to statistical learning*. Vol. 112. Springer.

Meredith, Marc (2004). “Why do universities compete in the ratings game? An empirical analysis of the effects of the US News and World Report college rankings”. In: *Research in Higher Education* 45.5, pp. 443–461.

Reiter, Andrew G. (2021). *U.S. News & World Report Historical Liberal Arts College and University Rankings*. URL: http://andyreiter.com/datasets/ (visited on 04/11/2021).

APPENDICES

A TABLES

Table A1

Descriptive Statistics

Variable	Mean	SD	C_V	Min.	Max.	Obs.
Dependent Variables						
Admission rate	0.31	0.18	57	0.04	0.80	898
SAT Average	1364.18	98.40	7	1067.00	1566.00	880
Undergraduates	14071.96	11033.91	78	879.00	43709.00	1149
Graduates	8628.52	5861.37	68	1039.00	34852.00	1149
Tuitionfee in-state	28596.37	15822.53	55	1941.00	59430.00	940
Tuitionfee out-of-state	34828.03	10418.46	30	7812.00	59430.00	943
Control Variables						
A. Student characteristics						
SAT Reading median	671.38	50.58	8	530.00	770.00	872
SAT Maths median	699.66	44.31	6	590.00	795.00	872
PP: Computer & Information[a]	0.04	0.04	106	0.00	0.33	1149
PP: Engineering	0.14	0.13	97	0.00	0.68	1149
PP: Maths & Statistics	0.02	0.02	84	0.00	0.18	1149
PP: Social Sciences	0.18	0.09	51	0.00	0.41	1149
PP: Business & Management	0.09	0.09	94	0.00	0.33	1149
PP: History	0.04	0.03	76	0.00	0.17	1149
UDGS: White	0.49	0.14	29	0.13	0.80	461
UGDS: Black	0.05	0.02	40	0.01	0.11	461
UGDS: Hispanic	0.11	0.05	48	0.03	0.27	461
UGDS: Asian	0.18	0.09	53	0.03	0.51	461
Age at entry	20.68	1.09	5	18.83	24.55	1000
Female	0.52	0.07	14	0.22	0.66	1000
1^{st} Generation Student	0.20	0.08	40	0.08	0.44	973
Parents Middle School Ed.	0.02	0.02	85	0.00	0.10	850
Parents High School Ed.	0.18	0.06	35	0.07	0.40	850
Parents Postsecondary Ed.	0.80	0.08	10	0.56	0.92	973
B. University characteristics						
Control[b]	1.66	0.47	29	1.00	2.00	1150
Avg. Faculty Salary	11131.67	2780.43	25	6082.00	22924.00	1100
Fulltime Faculty	0.80	0.12	16	0.36	1.00	850
Retention rate	0.96	0.02	2	0.87	1.00	748
Completion rate 100	0.72	0.16	22	0.20	0.92	1043
Completion rate 150	0.87	0.07	8	0.64	0.98	1097
Compl. degree within 2 years	0.16	0.11	69	0.00	0.67	945
Compl. degree within 3 years	0.58	0.22	38	0.01	0.92	946
Compl. degree within 4 years	0.69	0.20	29	0.06	0.95	898
C. Cost, Finances & Aid						
Pell grant ever	0.42	0.15	35	0.18	0.81	1000
Pell grant current year	0.17	0.08	45	0.05	0.48	550
Federal loan current year	0.31	0.14	45	0.02	0.83	550
Avg. annual cost	48803.70	16736.55	34	15395.00	75735.00	500
Net tuition revenue	18148.74	10450.59	58	2636.00	155111.00	1049
Instructional expenditure	30922.26	23851.46	77	1407.00	201515.00	1048

[a] The *PP* tags indicate program percentages.

[b] Control of institution: 1 for public institutions, 2 for private institutions.

Table A2

University Cluster Affiliations

(a) Technical Universities

	University
1	California Institute of Technology
2	Georgia Institute of Technology
3	Massachusetts Institute of Technology
4	Rensselaer Polytechnic Institute
5	Carnegie Mellon University

(b) Elite Universities

	University
1	Yale University
2	University of Chicago
3	Northwestern University
4	Johns Hopkins University
5	Harvard University
6	Washington University in St Louis
7	Dartmouth College
8	Princeton University
9	Columbia University
10	Duke University
11	University of Pennsylvania
12	Brown University
13	Vanderbilt University
14	Rice University
15	Stanford University

(c) Second-best Universities

	University
1	University of Southern California
2	Georgetown University
3	Emory University
4	University of Notre Dame
5	Tulane University of Louisiana
6	Boston College
7	Brandeis University
8	Tufts University
9	University of Michigan-Ann Arbor
10	Cornell University
11	New York University
12	University of Rochester
13	Wake Forest University
14	Case Western Reserve University
15	Lehigh University
16	William & Mary
17	University of Virginia

(d) Mass Universities

	University
1	University of California-Berkeley
2	University of California-Davis
3	University of California-Irvine
4	University of California-Los Angeles
5	University of California-San Diego
6	University of California-Santa Barbara
7	University of Florida
8	University of Illinois
9	University of North Carolina
10	Pennsylvania State University
11	The University of Texas at Austin
12	University of Washington
13	University of Wisconsin-Madison

Table A3

Means of Control Variables by Cluster

	Tech (N=5)	Elite (N=15)	Second-best (N=17)	Mass (N=13)
A. Student characteristics				
SAT Reading median	689.33	719.45	670.23	609.34
SAT Maths median	739.99	736.00	692.51	650.21
PP: Computer & Information	0.12	0.03	0.03	0.03
PP: Engineering	0.45	0.12	0.10	0.09
PP: Maths & Statistics	0.05	0.03	0.02	0.02
PP: Social Sciences	0.03	0.23	0.18	0.16
PP: Business & Management	0.07	0.03	0.15	0.10
PP: History	0.00	0.05	0.03	0.03
UGDS: White	0.44	0.46	0.56	0.45
UGDS: Black	0.05	0.07	0.06	0.04
UGDS: Hispanic	0.10	0.10	0.08	0.14
UGDS: Asian	0.24	0.18	0.12	0.22
Age at entry	19.83	21.04	20.48	20.85
Female	0.35	0.52	0.54	0.55
1^{st} Generation Student	0.15	0.15	0.16	0.30
Parents Middle School Ed.	0.01	0.02	0.01	0.04
Parents High School Ed.	0.15	0.14	0.15	0.26
Parents Postsecondary Ed.	0.85	0.85	0.84	0.70
B. University characteristics				
Rank	23.09	9.21	28.53	37.77
Control	1.80	2.00	1.82	1.00
Avg. faculty salary	11883.33	12465.32	10319.07	10366.35
Fulltime faculty	0.92	0.80	0.75	0.82
Retention rate	0.96	0.97	0.95	0.94
Completion rate 100	0.64	0.84	0.78	0.56
Completion rate 150	0.84	0.93	0.87	0.81
Compl. degree within 2 years	0.10	0.22	0.15	0.13
Compl. degree within 3 years	0.48	0.67	0.66	0.41
Compl. degree within 4 years	0.63	0.76	0.75	0.55
C. Cost, Finances & Aid				
Pell grant ever	0.41	0.36	0.35	0.58
Pell grant current year	0.14	0.14	0.14	0.26
Federal loan current year	0.33	0.19	0.38	0.37
Avg. annual cost	52034.42	60451.37	53867.69	27499.35
Net tuition revenue	17736.98	22777.34	20277.87	10189.96
Instructional expenditure	36473.28	52163.17	21684.16	16345.59

B FIGURES

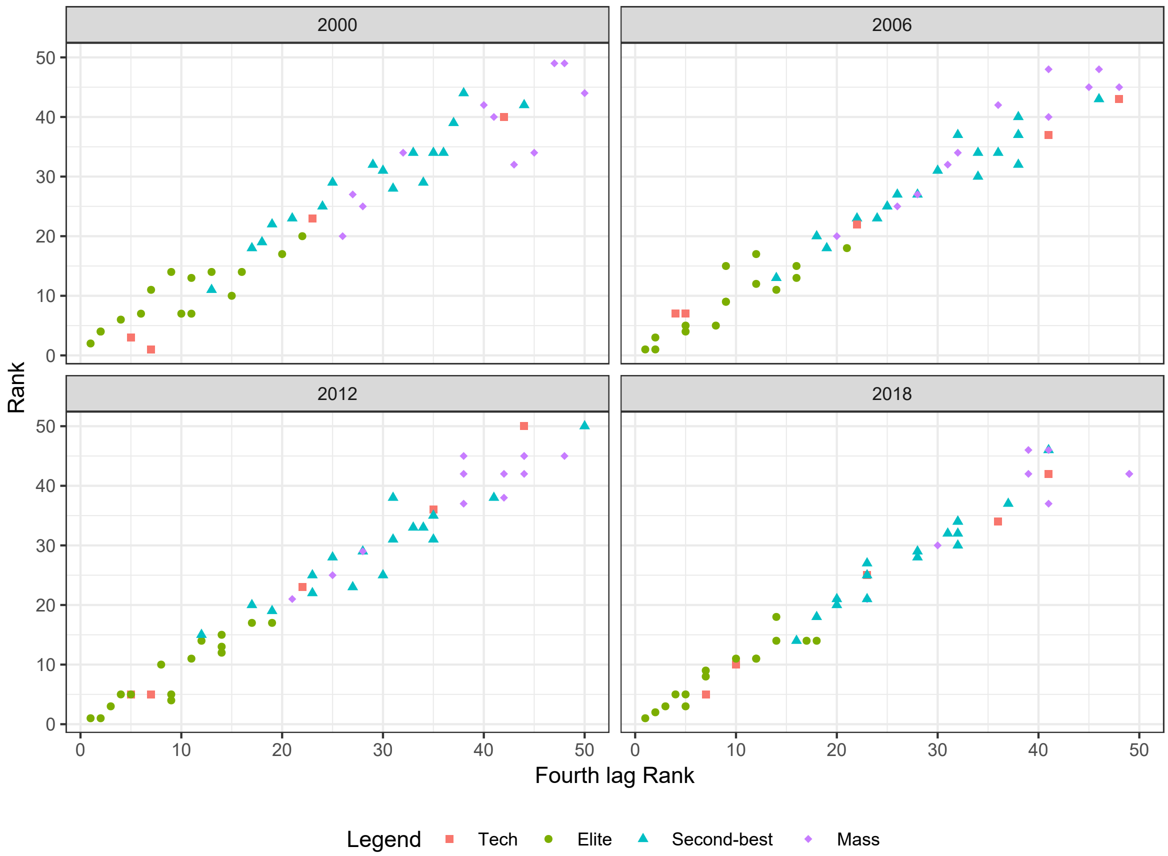

Figure B1: Correlations of Ranking Placements with Fourth Lags

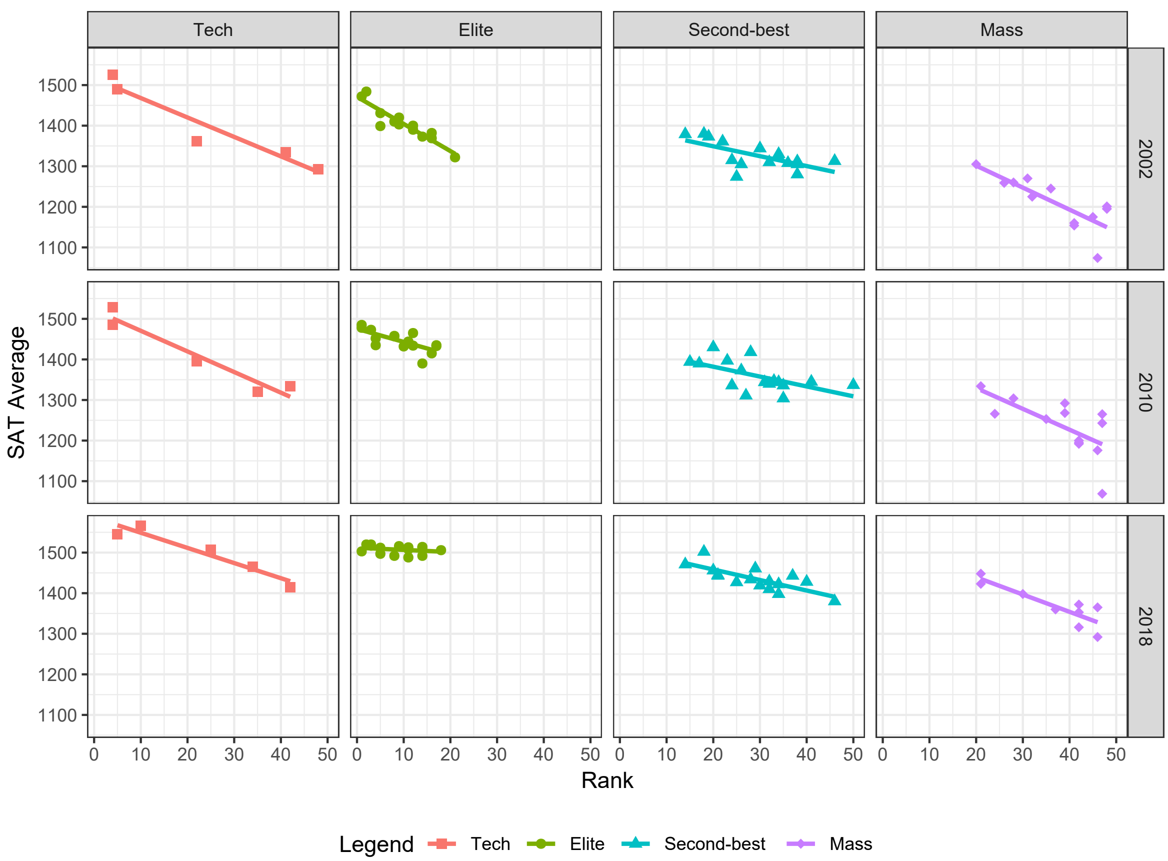

Figure B2: Correlation between SAT Average and Ranks by Cluster and Year

DECLARATION OF AUTHORSHIP

I assure that I wrote this seminar paper on my own without the usage of aids other than indicated. I only used the indicated sources. All literally or analogously translated paragraphs of the original sources are marked as such. The work has not been submitted to any examination office in the same or a similar form. Hereby I agree that the digital version of the work may be checked by plagiarism detection software.

C. Vennemann