DEMAND FOR EMPIRICAL COURSE DATA

An Undergraduate Research Scholars Thesis

by

JONATHAN A TILLINGHAST

Submitted to the Undergraduate Research Scholars program at
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by Research Advisor: Dr. Jason Lindo

May 2019

Major: Economics

TABLE OF CONTENTS

ABSTRACT

Demand for Empirical Course Data

Jonathan Tillinghast
Department of Economics
Texas A&M University

Research Advisor: Dr. Jason Lindo
Department of Economics
Texas A&M University

We perform a randomized control trial in which we provide students with a tool that visualizes empirical grade information of classes at a large public university. Visualization viewership is monitored and demand is linked to individuals through a unique student ID number. In this study, we test demand for such a course tool across numerous student groups. We find that freshmen are the most likely to use the visualization, quantitatively oriented majors are no more likely to use the visualization, and students are most likely to use tool shortly after receiving an email.

NOMENCLATURE

TAMU	Texas A&M University
UIN	University Identification Number
Viz	Visualization
RCT	Randomized Control Trial
GPA	Grade Point Average
ICD	International and Cultural Diversity Attribute
AG	College of Agriculture and Life Sciences
AR	College of Architecture
BA	College of Business
ED	College of Education and Human Development
EN	College of Engineering
GE	College of Geosciences
SC	College of Science
VM	College of Veterinary Medicine
OL	Liberal Arts majors other than Social Science
EC	Economics Major
PS	Political Science Major
PY	Psychology Major
SO	Sociology Major
GS	General Studies Major/College
U1	Freshman
U2	Sophomore
U3	Junior
U4	Senior

CHAPTER I

INTRODUCTION

Students consistently make course decisions lacking information about this process can be difficult and confusing particularly for freshman and students with small networks. This is because colleges rarely provide easily digestible information on previous course results. Students rely on a number of outlets to gather information about previous courses such as Rate My Professor, by word of mouth, or a number of other outlets. These can be very biased, unreliable, and of suspect usefulness. We are interested in testing consumer demand for empirical information on courses conveniently presented for comparison by students.

A similar study was previously done by Stanford University, where students were exposed students to similar grade information, but in a format that was not as good for comparison; multiple classes were not visible on the same screen but the tool was able to confer additional useful information such as weekly time commitment as judged by students who had taken the course (Chaturapruek et al. 2018). Additionally, our study does not allow the control group to view the tool, leaving them to use other, potentially less useful, course data visualizers available online which use the same source dataset.

Studies relevant to our discussion of the research include a 15-year long natural experiment conducted at Cornell. At the beginning of the 15-year period, Cornell began publicly publishing course grade information. They find when students were exposed to data on the mean grades of Economics courses, they tended towards the more leniently graded courses (Bar et al. 2009) and low ability students had a stronger tendency to enroll in leniently graded courses. This indicates to us that students do in fact demand information on grades and make decisions based off of that information. We are interested to see if this would have a different effect at a university where the students are of comparatively lower ability than Cornell.

Through completion of this project, we have contributed to the body of knowledge regarding the state of student demand for streamlined, reliable, comprehensive historical course information. The office of the registrar at Texas A&M already publicly provides access to grade distributions; however, it is difficult to compare across classes given they are in text format and not aggregated across years. We downloaded and aggregated the data in order to make the visualization. Other websites have downloaded the same data from Texas A&M's registrar and made other course selection tools. Some offer a graphic schedule of what a week would look like given registration in certain classes, while others provide information aggregated over time in text form. We have developed a tool to make this grade information accessible for comparison and simple to interpret – in contrast to the other tools where it is hard to make a side by side comparison of which class offers a historically higher GPA.

Randomizing across different groups of credit hours earned denoted by the university as U1, U2, U3, and U4, we sent emails to 5,095 students strategically with the link to the course tool a week before general registration opened and to each class year three days before their registration time. Graduate students were not included in the study. A screen shot of the email sent is provided in the Appendix. It is important to note that most of studies previously performed regarding grade information and related student outcomes were performed at elite universities. Texas A&M is more representative of US universities as a whole, thus its results lend themselves to greater external validity.

CHAPTER II

METHODS

In this project we employ a randomized control trial (RCT) research design. We expose students to empirical grade information including course name, department, number, professor average course grade (0.0-4.0), percentages of each grade type (A, B, C, D, F, Q), core curriculum, honors, and international and cultural diversity (ICD) attributes. Historical course information summaries are derived using data on courses and professors from the past 17 semesters. All students must have certain amount of core requirement and ICD classes in order to graduate. Honors students need to take several honors attribute classes in order to complete their honors curriculum and graduate with distinction as honors student. Latin Honors (cum laude, etc.) are strictly based on GPA. Figure 1 is a screenshot of the visualization.

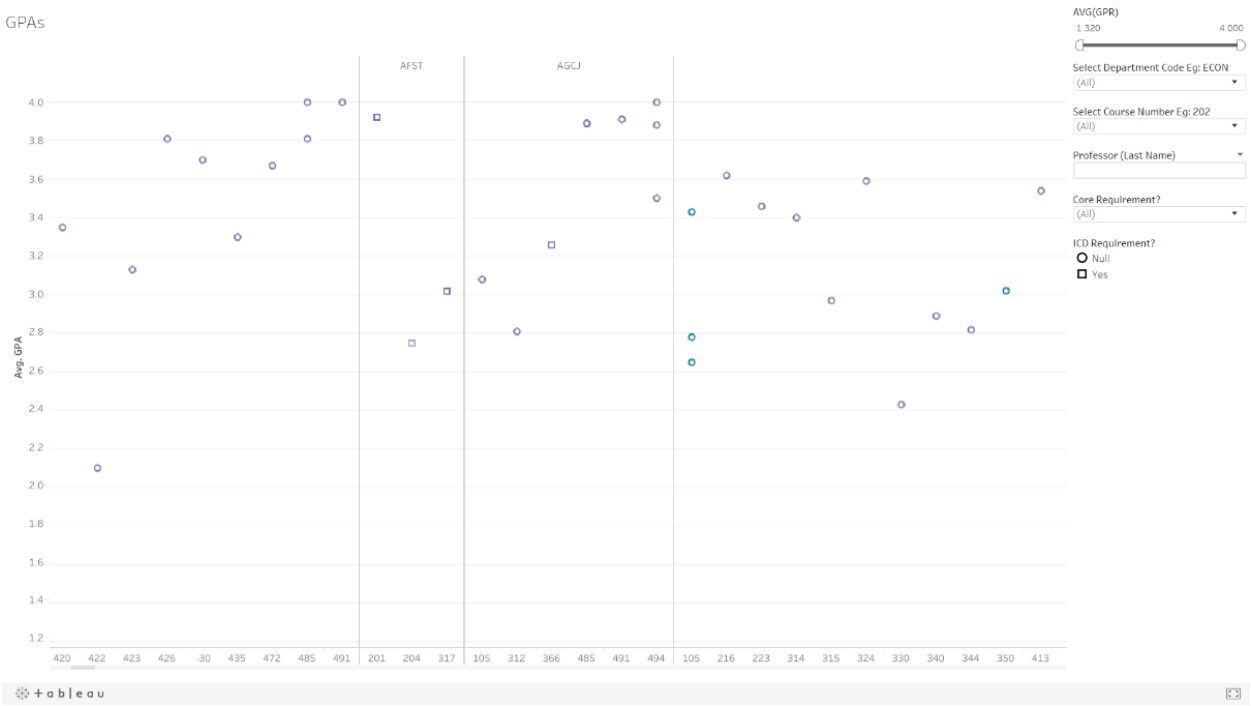

Figure 1. Course Tool

To conduct our RCT we sent out emails with a link to a website containing the course tool and a sign-up form to authenticate that they were in fact Texas A&M students using their "@tamu.edu" email address. In the process of authentication with a Texas A&M University email address, we also requested their academic major. Emails were sent to students based on subgroups of the student population by university class level ranging from freshmen with zero credit hours to seniors. U1-U4 students are freshman (U1) through seniors (U4). The different groups received an email at a different time based upon their predetermined registration schedule. We sent an email to all students in the sample a week before general registration opened and three days before each group's registration time so they would use the course tool at a time when registration is a priority for them and still have time to alter what they plan to take.

CHAPTER III

RESULTS

At the beginning of the Spring 2019 semester, after students were able to add or drop classes, we collected our results and shut down the website servicing the course tool. A total of 5,095 students received an email, of which 997 signed in. When looking at student usage per day, students were most likely to use the course tool shortly after they received an email about the course tool - the spikes in the data were all shortly after an email was sent. See figure 2 below.

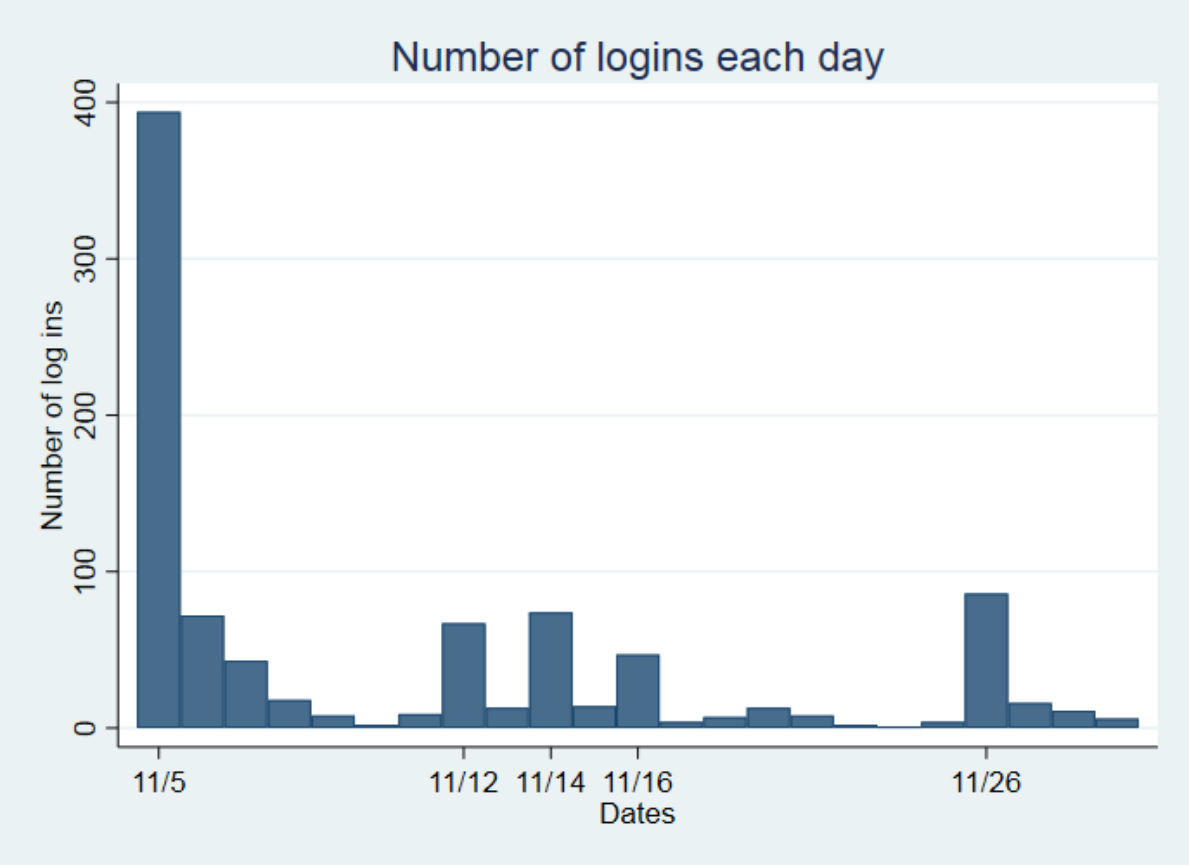

Figure 2. Logins

In the following table we provide descriptive statistics for a breakdown of students who received an email and logged in by class year and by college. The Liberal Arts college consists of many different types of

majors we split this college into different sections; Economics, Political Science, Psychology, Sociology, and Other. Some interesting results to note are that freshmen are had the highest rates of logging in which is exactly what we expected. When looking at login results by college or major, we find that Economics students had the highest rates of logging in. A high take up by Economics students is in line with how we would expect them to act. However, this result could partially be attributed to in group preferences due to the sender noting he was a PhD student in the Economics department. The college with the second highest logins was Veterinary Medicine/ Biomedical Sciences. In table 1 below we provide regression results for the causal impact on take up.

Table 1. Descriptive statistics

Class Year	Total Students	Received Email	# Logged In	% Logged In
U1	10,637	1,064	277	26.00%
U2	11,336	1,134	215	19.00%
U3	11,422	1,142	225	19.70%
U4	17,079	1,708	219	12.80%
College	Total Students	Received Email	# Logged In	% Logged In
AG	6,425	635	105	16.50%
AR	2,669	245	23	9.40%
BA	5,110	512	76	14.80%
EC	940	99	29	29.30%
ED	5,185	507	84	16.60%
EN	15,355	1551	305	19.70%
GE	1,012	109	23	21.10%
GS	2,205	211	31	14.70%
OL	3,555	338	77	22.80%
PS	996	110	23	20.90%
PY	1,528	164	33	20.10%
SC	2,616	274	60	21.90%
SO	547	60	9	15.00%
VM	2,653	266	62	23.30%

To investigate the effects of Course Tool Usage we used several regression models. The first of which is defined as follows:

$$Logged\_in_i = \alpha + \beta Received\_Email_i + \varepsilon_i$$

for each $i \in Uk$ where $k \in \{1,2,3,4\}$

We defined logged_in as a dummy that is 1 if anybody logged in to use the course tool and a 0 if not. Received_email is a dummy that is 1 if somebody received an email and 0 if not. This analysis is done conditional on each class year. We present these regression results below:

Table 2. Impact of Receiving Email on Course Tool Usage by Class Year

Variables	U1	U2	U3	U4
1=received email, 0=didn't	0.251***	0.176***	0.192***	0.125***
	(0.005)	(0.004)	(0.004)	(0.003)
Observations	10,637	11,336	11,422	17,079
R-squared	0.224	0.150	0.172	0.111

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Results represent OLS regressions

Looking at the results above, we find that indeed, freshmen do have the highest rates of usage at about 25%. U2's and U3's have slightly lower rates at 17.6% and 19.2% respectively. Seniors have the lowest rates of usage at 12.5%. These results are consistent with our prior beliefs that underclassmen will have higher usage rates. In Figure 3 on the following page we show these results visualized in an efficient manner:

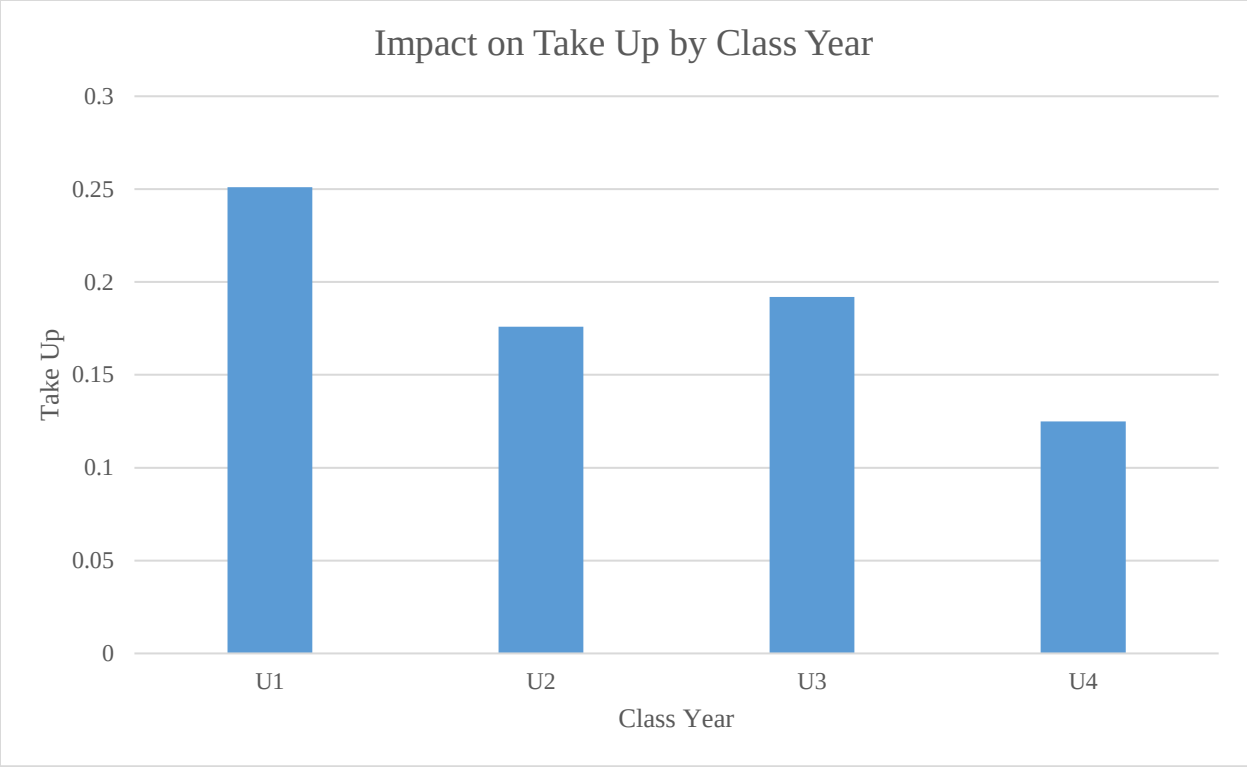

Figure 3. Take Up by Year

To understand the full range effects of our course tool, we additionally break down the impacts by college/major. We estimate the following regression conditional on each class year:

$$logged\_in_i = \alpha + \sum_{j \in J} \delta_j College_{ji} + \sum_{j \in J} \beta_j [College_{ij} * Received\_Email_i] + \varepsilon_i$$

$$\text{for } i \in \text{U}k \text{ where } k \in \{1,2,3,4\}$$

Where J is the set of all colleges. The betas represent the impact of receiving the email on likelihood of using the tool within each college/major conditional on a student's class year.

Below we show our results from these regressions and it is shown that individuals with certain majors and colleges are substantially more likely to use the course tool than others. We make a distinction to separate

the Liberal Arts into different majors because there is a large degree of heterogeneity in the types of majors and people taking such majors within the college of liberal arts. For example: English majors do not belong in the same category as political science majors thus, we draw a distinction between the majors.

Some interesting results to note are that Geosciences freshmen had the highest rate of usage while General Studies freshmen had the lowest. For sophomores, Psychology majors had the highest usage rate compared to Architecture having the lowest. For juniors, Economics students had the highest usage rates versus again Architecture having the lowest. Finally, for seniors, Economics students again had the highest usage rates while Architecture continued to have the lowest usage rates see Table 3 and Figure 4. These estimates provide clear evidence that our emails had a statistically significant impact on students that vary based on colleges/majors and class year. Furthermore, in Figure 3 we display these results in a graphical manner that show compelling visual evidence. Another interesting results to note is the trend of usage being highest for freshmen, dipping for sophomores, jumping again for juniors, then dropping for seniors with a college/major.

Table 3. Impact of Receiving Email on Course Tool Usage by College

VARIABLES	U1	U2	U3	U4
AG	0.266***	0.155***	0.196***	0.098***
	(0.015)	(0.011)	(0.011)	(0.007)
AR	0.188***	0.045**	0.108***	0.077***
	(0.026)	(0.020)	(0.016)	(0.011)
BA	0.196***	0.175***	0.136***	0.097***
	(0.015)	(0.013)	(0.011)	(0.009)
EC	0.300***	0.244***	0.340***	0.172***
	(0.033)	(0.030)	(0.025)	(0.021)
ED	0.361***	0.135***	0.178***	0.083***
	(0.017)	(0.013)	(0.012)	(0.008)
EN	0.243***	0.186***	0.200***	0.144***
	(0.007)	(0.007)	(0.008)	(0.005)
GE	0.364***	0.143***	0.296***	0.163***
	(0.044)	(0.029)	(0.025)	(0.016)
GS	0.175***	0.135***	-0.000	
	(0.015)	(0.013)	(0.038)	
OL	0.267***	0.268***	0.226***	0.146***
	(0.022)	(0.014)	(0.013)	(0.011)
PS	0.263***	0.190***	0.318***	0.094***
	(0.034)	(0.023)	(0.028)	(0.020)
PY	0.179***	0.276***	0.200***	0.136***
	(0.032)	(0.024)	(0.018)	(0.015)
SC	0.262***	0.169***	0.232***	0.202***
	(0.016)	(0.017)	(0.019)	(0.012)
SO	-0.000	0.250***	-0.007	0.128***
	(0.142)	(0.038)	(0.046)	(0.018)
VM	0.343***	0.206***	0.200***	0.163***
	(0.018)	(0.017)	(0.017)	(0.013)
Observations	10,637	11,336	11,422	17,079
R-squared	0.234	0.162	0.187	0.120

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Results represent OLS regressions

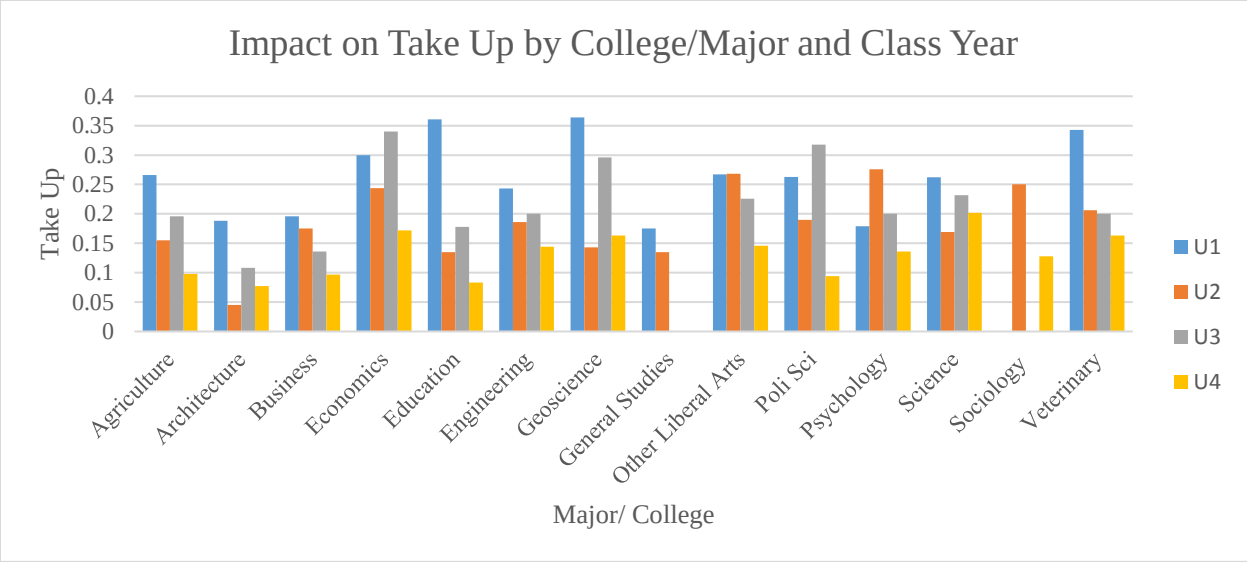

Figure 4. Take Up by College

Our results closely mirror the Stanford study with respect to course tool usage by class year. In the Stanford study, seniors were the least likely to use their course tool, Carta, while freshman were the most likely to use the tool (Chaturapruek et al. 2018). Because seniors are less likely to use a course tool than freshman at multiple types of institutions, elite and nonelite, our results gain some external validity.

CHAPTER IV

CONCLUSION

Our project is the first of its kind to be performed at a large public research university. Many more students attend these types of universities so it is important to understand how these students respond to empirical grade information. Additionally, our course tool displays empirical course information in a way that provides a higher degree of comparison capability than others previously available. Information is known to have effects on student decision making (Bar et al. 2009, Ost et al. 2010). As such, it is important that we understand how students respond to exposure to such information such that university administrators can make proper decisions regarding what sorts information should be available to students. Should further research show that exposure to students respond negatively to having such information it would be in the best interest of the students to restrict access to the information.

We find that the information disparity between freshman and upperclassmen is reflected in their demand for information and students not seeking a degree are less likely to use the course tool. Most students use the course tool shortly after they received their emails. Quantitatively oriented majors are not necessarily more likely to use the course tool, despite the fact they are more skilled with numbers.

REFERENCES

Bar, T., Kadiyali, V., & Zussman, A.. (2009). Grade Information and Grade Inflation: The Cornell Experiment. Journal, 93-108.

Chaturapruek, S., & S, T.. (n.d.). How a Data-driven Course Planning Tool Affects College Students GPA. Proceedings, 18, 2018.

Grant, D.. (2007). Grades as Information. Economics, 201-14.

Lindo, J., Sanders, N., & Oreopoulos, P.. (n.d.). Ability, Gender, and Performance Standards: Evidence from Academic Probation.

Main, J. B., & Ost, B.. (2014). The Impact of Letter Grades on Student Effort, Course Selection, and Major Choice: A Regression-Discontinuity Analysis. The, 1-10.

Ost, B.. (2010). The Role of Peers and Grades in Determining Major Persistence in the Sciences. Economics, 923-34.

APPENDIX

Appendix 1. Email Screenshot

Howdy!

You have been selected to be given access to a previously unavailable course selection tool, which allows you to filter courses based on the requirements they satisfy, department, and historical grade distributions. If you would like to use this tool, the following link:

Access Course Information Tool (option 1) (option 2)

Thanks and Gig'em
David Pritchard
PhD Student
Department of Economics
Texas A&M University

You are receiving this email as a part of a study of student interest in this sort of tool