

some misc comments:

* format references as suggested in my comments on p 6 below

* make sure there are at least 5 references (and that each is cited in the body of the paper)

Undergraduate Prospects after Graduation

Introduction

Given the recent scandal revealing the over optimistic prospects for graduating law school students, the statistics produced by universities and published in the US News and World Report are being brought into question. These misleading statistics encourage hopeful JD seekers to pursue startling loans with the expectation that their debts will be paid off with relative ease upon graduation thanks to the supposed 84% job placement ratings.

While the production of undergraduate college rankings has often been criticized for its accuracy in measuring the actual quality of education, Carnegie Mellon University and other universities have long bolstered their reputations for producing intelligent, motivated, and successful students with the use of these faulty lists. However, this raises the question of how measurably successful Carnegie Mellon University undergraduate alumni are. Where do alumni relocate to? What occupations do they practice? What graduate programs do they choose to pursue? And, according to the common man's perception of comparative success, how do Carnegie Mellon University undergraduate alumni stack up when compared to graduates of other universities?

Others have conducted analyses of the correlation between academic quality and university ranking systems such as David D. Dill's and Maarja Soo's "Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems" published by "Higher Education," "International University Ranking Systems and the Idea of University Excellence" by Paul Taylor and Richard Braddock from the "Journal of Higher Education Policy and Management," and *College Rankings Exposed: the Art of Getting a Quality Education in the 21st Century* by Paul Boyer. Each of these articles and the book discusses the concept of ranking systems as a measure of excellence, but no study has directly applied these concepts of ranking as excellence to Carnegie Mellon University.

This study analyzes the data collected and evaluated by the Carnegie Mellon University Career Center (http://www.studentaffairs.cmu.edu/career/students_alumni/post-grad-survey/index.html) in order to answer such questions as: Are alumni—successful by Carnegie Mellon standards—well received by employers and graduate programs? Do alumni display a tendency to remain near to Pittsburgh or to relocate elsewhere? Do alumni successfully attain employment relevant to their subject(s) of study? How accurately do national and international rankings systems gauge the value of a Carnegie Mellon Undergraduate degree?

The data collected from the Carnegie Mellon University Career Center is will also be used to test the effectiveness of various survey methods. The Career Center has nearly perfected its collection data collection methods as response rates generally run somewhere in the 90th percentile (with the exception of College of Fine Arts classes where response rates are as low as the 70th percentile). This has yielded results near to that of census data. As such, this study assesses the effectiveness of certain types of sampling schemes (stratified and clustered sampling) to produce results representative of the target population so that future statistical researchers can visually apprehend the significance of various survey designs. Few have had the data provided or the opportunity to conduct a study on accurate census data in order to optimize survey results to population parameters.

Methods

Our target population was graduating Carnegie Mellon undergraduates of the Pittsburgh campus from the years 2005 to 2010. The sampling frame that came with this data originated from the career center's post graduation survey. This data from the Career Center consistently includes greater than 94%-71% response rates. Majors within CFA had the lowest set of response rates. This may be due to the significantly smaller class sizes, the personalities of the students, and the difficulties associated with categorizing employment in the arts. With the administrative records from the career center we attempted to reconstruct the entire graduating classes from 2005-2010. Each person who answered the survey and either was employed or went on to graduate school was listed individually. For the employed graduates the company name, job title, and location of job by city and state were given. At times these jobs were not in the United States and therefore had a name and country, a country, or just a city. Those who went on to graduate school were listed by university, by specific program, and then whether they were getting a Masters or PhD degree. From these individual listings we were able to create a spreadsheet that contained each individual from the post-graduation survey.

Our sample design consisted of two very different approaches. Based on the wealth of information provided by the Career Center we were able to construct a sampling frame that was almost exactly the same as the target population. This allowed us to conduct tests on essentially the data from the whole population, which would make our study a census. We then did a second round of analysis where we used various types of sampling methods. The first time did a simple random sample of the students in each year. We took 2005 through 2010 and took simple random samples each one. This was then combined to calculate all of our inference tests. In each year we took various sizes of samples. Based on our calculations that come later we found the ideal sample size per year to be 292 students. In a discussion with Professor Junker we examined the idea of using various sizes of simple random samples and compare them to the census data we had. Our sample sizes ended up being 100, 300, 600, and 900 students from a single year. We then stratified the data and broke all of the students into the six colleges that are part of the university: the College of Fine Arts (CFA), Carnegie Institute of Technology (CIT), the College of Humanities and Social Sciences (HSS), the Mellon College of Science (MCS), the School of Computer Science (SCS), and the Tepper School of Business (Tepper). From there we stratified the data again and took each department in the school and pulled a random sample from each of those. Our next method involved stratifying the data by college again and then using each department as a cluster. From there we randomly picked on cluster or department in each school and sampled every single student in those clusters. Based on the clustered and stratified by department example we will test these results with the census data to see how accurate these prediction methods are for our study.

Sample Size Calculations

Without Replacement

$$n > (N * n_0) / (N + n_0) \quad n_0 = (z^2 (SD)^2) / (ME)^2$$
$$z = 1.96$$

SD = 0.5
ME = .05
N = 1200 (estimate)

$$n_0 = ((1.96)^2(0.5)^2)/(.05)^2 = 384.16$$

n = 292

?

We do not have any sort of data on the variance of the two colleges when it comes to our most important question: the percentage of Carnegie Mellon Undergraduates who gain employment in a workplace that utilizes the knowledge and skills they learned as a result of their particular major, it is very difficult to calculate any kind of sample size estimate using stratification. Unless the groups are significantly different there will be no gain from stratifying in the sample size calculations.

We used the data of about 7200 individual students. About each year 1200 students graduate and we have 6 years worth of responses. Our nonresponse rate was based purely on the ability of the Career Center to collect responses from individuals. As stated previously the numbers that they provide are fairly good. We will attempt to use imputation when we run into individuals that do not respond. This will make it so when we actually select our samples from the frame there will be no nonresponse. Our information is limited to what the Career Center will give us. In order to increase the variables that we have we are matching the information we have to external databases. We will be using www.salary.com to add in salary information based on the job title. With the aid of USNews and Fiske we will get information about the rankings of college graduate programs. Some of the summary sheets from the Career center have aggregate information about salaries so we will use that to check our additional data from salary.com.

Since we are looking at administrative records our questionnaire is a list of variables that we will look for when examining an individual. The primary ones include salary in present dollars, the distance they end up from Pittsburgh, PA, whether the current job they have requires or uses the knowledge and skills learned in undergraduate study, the ranking of the graduate program relative to Carnegie Mellon's ranking in that area, and whether the student is employed.

Our post survey processing is the entering of the data from the Career Center sheets to an excel file. We will have multiple people go through to check whether they think that the job description matches the major. Since the titles can be very ambiguous we were hoping to look at the ones where we did not agree.

Results

We will conduct several ANOVA tests. First we will be looking at the mean salaries across the majors for our census data. We will then do ANOVA testing between the various sampling techniques.

We will do binomial tests for the proportion questions, such as whether a person went to graduate school or not.

We will calculate the descriptive statistics for many of our variables as well.

great

Discussion

For this survey, we are trying to show CMU undergraduate prospects after they graduate so students can exactly know what awaits them once they graduate. We did not design the actual survey. There is an annual survey that Carnegie Mellon Career Center conducts on undergraduates who are graduating Carnegie Mellon University. The purpose of Career Center's survey is to keep record of undergraduate prospects after graduation such as where they are employed, their job title, salary, and whether they are going to graduate school. Our project was to take this data and analyze it further using different survey methods while adding our own variables. These new variables include correlation between one's job title and his major, the distance of one's employment place from Pittsburgh, whether the graduate school a student has enrolled is more/less prestigious than that of Carnegie Mellon using US News Ranking and Fiske for Fine Arts school.

Some of the strengths of this project are very high response rate and the fact that we were able to save lot of time by skipping the part where we actually conduct the survey meaning that we are able to allot more time to analysis of the data. Because we used data from Career Center, we were able to acquire data with response rate of over 90% in most cases with an exception of couple 70%'s, which is still considered very high. Because Career Center conducts this survey on every undergraduate student who is graduating, we can be confident that there is very little sampling error.

However, there are also many critical flaws in our project that stems from the fact that we used Career Center's survey. One of them is that because we are not the one who designed the actual survey we are restricted to what kind of analysis we can do; we cannot tweak the survey questions to get better data that will suit our analysis purpose better. Second flaw is that Career Center did not disclose all the data to us. We do not know the exact individual's salary; we only know the min, max, and average of students' salaries grouped by year and major. As a result, we decided to get approximate salary amount from salary.com but given that the many job titles are ambiguous, it is inevitable that we will have much error.

We are still in process of compiling the data from Career Center into Excel for more efficient analysis. Once data compilation is done we will start our analysis and will be able to find out if there are relationships between various variables and whether the employment rate that CMU has provided us is accurate and true.

Appendix

Great start. Please adjust formatting to agree with the examples in the 2nd half of <http://www.library.cornell.edu/resrch/citmanage/apa>

Also, list of references should come before appendices

List of references

1. Taylor, Paul, and Richard Braddock. "International University Ranking Systems and the Idea of University Excellence." *Journal of Higher Education Policy and Management* 29.3 (2007): 245-60. Print.
2. Dill, David D., and Maarja Soo. "Academic Quality, League Tables, and Public Policy: A Cross-national Analysis of University Ranking Systems." *Higher Education* 49.4 (2005): 495-533. Print.
3. Boyer, Paul. *College Rankings Exposed: the Art of Getting a Quality Education in the 21st Century*. 2004. Print.

Questionnaire

The "Analysis of Carnegie Mellon Undergraduate Prospects After Graduation Statistics" survey that Group H proposes wishes to measure variables including:

It would be useful to include summary statistics with each question below in the appx of your final report

1. percentage of CMU alumni who find employment upon graduation (for whole sample and each of five sample years)
 - a. this variable would be calculated using the Career Center data on salary as a reference
2. percentage of CMU alumni from each college finding employment upon graduation (for whole sample and each of five sample years)
 - a. this variable would be calculated using the Career Center data on employment percentages as a reference
3. percentage of CMU alumni from each major finding employment upon graduation (for whole sample and each of five sample years)
 - a. this variable would be calculated using the Career Center data on employment percentages as a reference
4. percentage of CMU alumni whose employment relates to their major (for whole sample and each of five sample years)
 - a. this variable would be calculated using the Career Center data on employer and job title
 - b. both employer and job title of each alumnus will be coded as "does relate to major" or "does not relate to major"
 - i. for a job to be coded as "does relate to major" the job title should indicate that the undergraduate degree received is necessary to fulfill some aspect of the job qualifications
 1. e.g. the job title "Mechanical Designer Engineer" *does relate* to a major in engineering/physics

- ii. coding will be conducted by several Group H members to ensure that coding is consistent by person
- 5. percentages of the job locations (states) of CMU alumni
 - a. this variable would be calculated using the Career Center data on job location
- 6. percentage of CMU alumni enrolling in graduate school
 - a. this variable would be calculated using the Career Center data on the number of alumnus pursuing graduate degrees in respect to his/her class
- 7. average comparative ranking of graduate school program compared to CMU ranking
 - a. this variable would be a measurement of the deviation from CMU ranking for each graduate program pursued by CMU alumni
 - b. CMU ranking would be the reference point equal to zero
 - c. graduate program ranking would be measured as a negative or positive number in comparison with CMU ranking
 - i. graduate program ranking would be collected from *US News and World Report*
 - ii. negative graduate program rankings would represent programs of lesser prestige than CMU
 - iii. positive graduate program rankings would represent programs of greater prestige than CMU
- 8. majors (top 5) from which graduates matriculate in the most prestigious graduate schools
 - a. this variable would summarize the findings of variable (7) found above
 - b. for each major, the average comparative ranking will be calculated, and the majors with the highest average comparative rankings will be reported (top 5)
- 9. percentage of graduates pursuing graduate degrees at Carnegie Mellon
 - a. this variable would be calculated using Career Center data on graduate institution at which an alumnus is pursuing a graduate degree
- 10. most common graduate degrees (top 5) that alumni pursuing graduate degrees at Carnegie Mellon are pursuing
 - a. this variable is similar to variable (8) found above
 - b. the numbers of alumnus pursuing individual graduate degree types would be compiled and the most popular (top 5) graduate degrees would be reported
- 11. mean real (without inflation) starting salary of graduates from CMU (for whole sample and each of five sample years)
 - a. this variable would employ concepts from macroeconomics and use the aggregate data on salary supplied by the Career Center
 - b. for each mean salary calculation:
 - i. mean nominal salary: mean salary from the sample year of concern
 - ii. 2011 will always be used as the reference for calculating the real salary
- 12. mean real (without inflation) starting salary of graduates from each college (for whole sample and each of five sample years)

- i.mean nominal salary: mean salary from the sample year of concern
 - ii.2011 will always be used as the reference for calculating the real salary
- 13. mean real (without inflation) starting salary of graduates from each major (for whole sample and each of five sample years)
 - i.mean nominal salary: mean salary from the sample year of concern
 - ii.2011 will always be used as the reference for calculating the real salary
- 14. deviation from the mean real (without inflation) starting salary of graduates from CMU in 2005
 - a. this variable should measure whether or not the mean real starting salary of graduates from CMU is decreasing or increasing
 - i.mean nominal salary: mean salary from the sample year of concern
 - ii.2011 will always be used as the reference for calculating the real salary
- 15. percentage of graduates electing to remain in Pittsburgh for employment or educational purposes
 - a. percentage of graduates electing to remain in Pittsburgh for employment
 - b. percentage of graduates electing to remain in Pittsburgh for education
 - c. this variable would be calculated using the Career Center's data on job location/graduate institution location
 - d. this variable could help indicate how undergraduates found their experience in Pittsburgh

Data From Career Center