The Trillion Dollar Problem For Tomorrow:

Use of Web Scraping to Explore Return of Investment of Higher Education

JON HARRIS

NYC DATA SCIENCE ACADEMY

Background

As of 2019, 45mill borrowers owe $1.5T in federal student loan debt [1]

College tuition is at an all-time high[2]

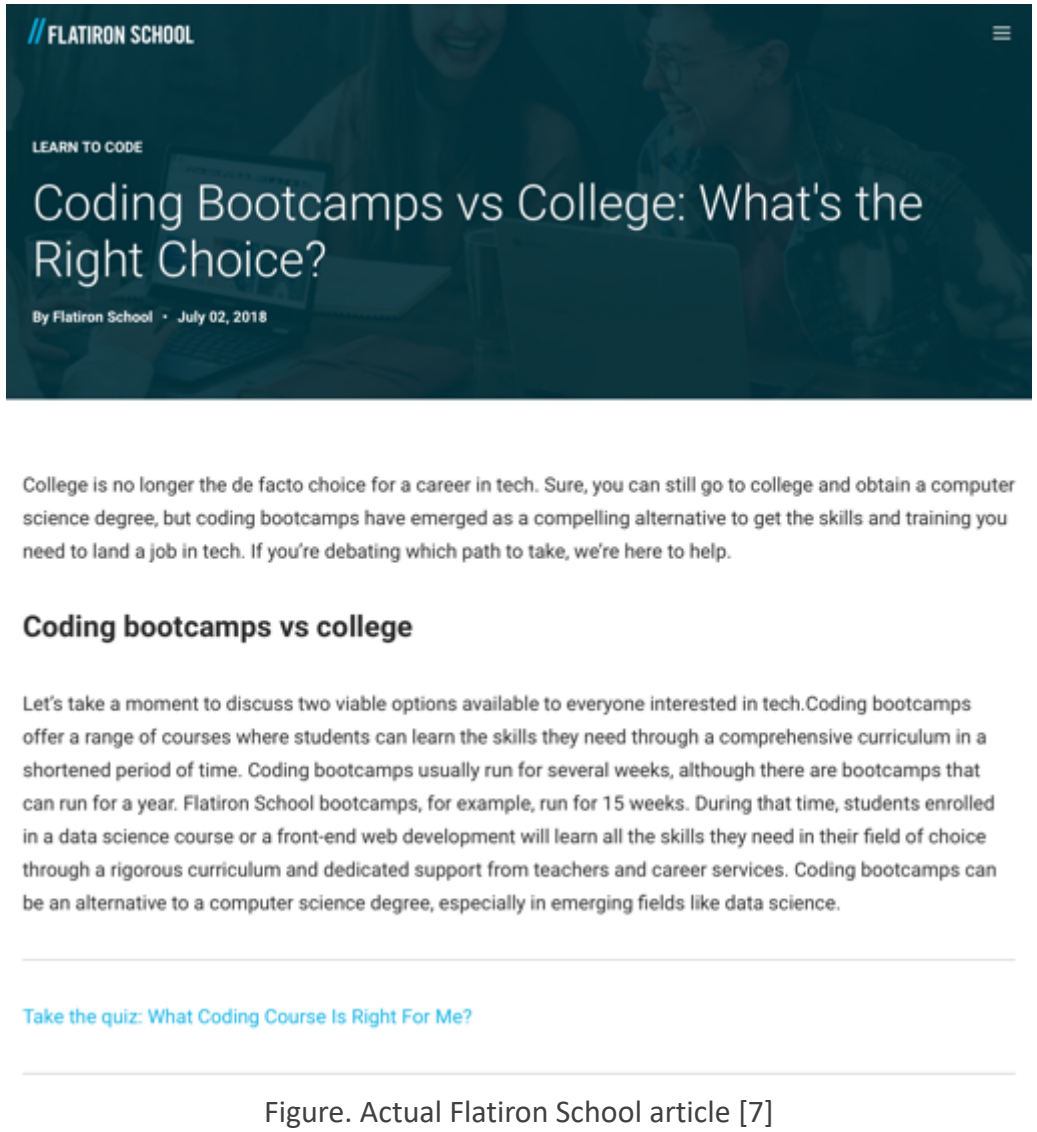
- 2018 avg. yearly tuition: Public, $21.4k; Private, $48.5k
- 1988 adj. avg. yearly tuition: Public, $9.4k; Private, $24.8k
  - Fun Fact: avg. hourly salary of post-Bachelor grads has increased 2.3% since 1989 [3]

University of California recently proposed a 5yr 2.8% annual increase [4]

- Top 50 universities have increased tuition by 3.6% [5]

Tuition hike largely attributed to loss in state funding [6]

- 2018: -13% across all states
  - AL, AZ, LA, MS, OK, and PA down between 30-60%



FLATIRON SCHOOL

LEARN TO CODE

Coding Bootcamps vs College: What's the Right Choice?

By Flatiron School • July 02, 2018

College is no longer the de facto choice for a career in tech. Sure, you can still go to college and obtain a computer science degree, but coding bootcamps have emerged as a compelling alternative to get the skills and training you need to land a job in tech. If you're debating which path to take, we're here to help.

Coding bootcamps vs college

Let's take a moment to discuss two viable options available to everyone interested in tech.Coding bootcamps offer a range of courses where students can learn the skills they need through a comprehensive curriculum in a shortened period of time. Coding bootcamps usually run for several weeks, although there are bootcamps that can run for a year. Flatiron School bootcamps, for example, run for 15 weeks. During that time, students enrolled in a data science course or a front-end web development will learn all the skills they need in their field of choice through a rigorous curriculum and dedicated support from teachers and career services. Coding bootcamps can be an alternative to a computer science degree, especially in emerging fields like data science.

Take the quiz: What Coding Course Is Right For Me?

Figure. Actual Flatiron School article [7]

Changing Educational Landscape

- Rising tuition costs changing how we view college
  - No longer for exploration/expanding one's mind
  - Forced pragmatism; college now considered an investment
- Mounting costs have led to new industries proposing alternatives to a 4yr college experience
  - Coding boot camps
  - Associate degrees
  - Job-specific apprenticeship programs
- HS grads must weight many decisions
  - Public schools
  - Private schools
  - College major
  - Financial aid

Online Resources For Decision Making

PAYSCALE.COM

Public-facing Research Center

- Education and salary data collected from 3.5 million civilian employees

Provides in-depth college and career-specific analytic reports for paying members

- Available data is stratified across each major, college, career field, and job

Free data is available to the public

- Includes general average salary and estimated return on investment data across universities and grouped college major types

Purpose

- Broadly inspired by my own nomadic college experience
  - Texas Tech → Drexel University → UPenn
- To explore how decisions made by a HS senior affect the return on investment of higher education
  - Type of degree
  - Type of college
  - College major
  - Regionality

HYPOTHESIS

1. Bachelors degree maximizes earning potential compared to a 2yr Associates program
2. Return on investment higher for 4yr private universities compared to public schools (but may be regional)
3. Return on investment are dissimilar between majors

Web Scraping Methodology

Websites used: 2018 College Salary Report

- Early- and Mid-career salary by state [8]
- Estimated ROI by state [9]
- Estimated ROI by major [10]

Web Scraping Tools:

- Python/Selenium
- Jupyter Notebook

Visualization Packages:

- Seaborn
- Matplotlib
- Plotly

Analytical Packages:

- Scipy
  - Pearson correlation, independent T-Test, One way Independent ANOVA
- Statsmodels
  - Tukey's post hoc analyses

CUMULATIVE VARIABLES SCRAPED FROM WEBSITES:

- College name
- College type
  - Private
  - Public (in state, out of state)
  - Service academy
- Early career salary
  - Median salary of alumni (0-5yrs)
- Mid-career salary
  - Median salary of alumni (5-10yrs)
- Estimated 20yr net ROI
  - Factors in 20yrs of median salary earned by bachelor-only graduate, 24yrs avg HS salary, and cost of school
  - With and without financial aid/grants
- %STEM Degrees
  - Percent degrees awarded in 2018
- Total 4yr cost
  - Integrated Postsecondary Education Data System (IPEDS)
- Typical years to graduate
  - Avg years required to graduate at least 65% to receive degree (IPEDS)
- Average loan amount

Results

Salary Metrics Per Degree Type

Early-Career Salary:

- Bachelors degree >17%

Mid-Career Salary:

- Bachelors degree > 37%

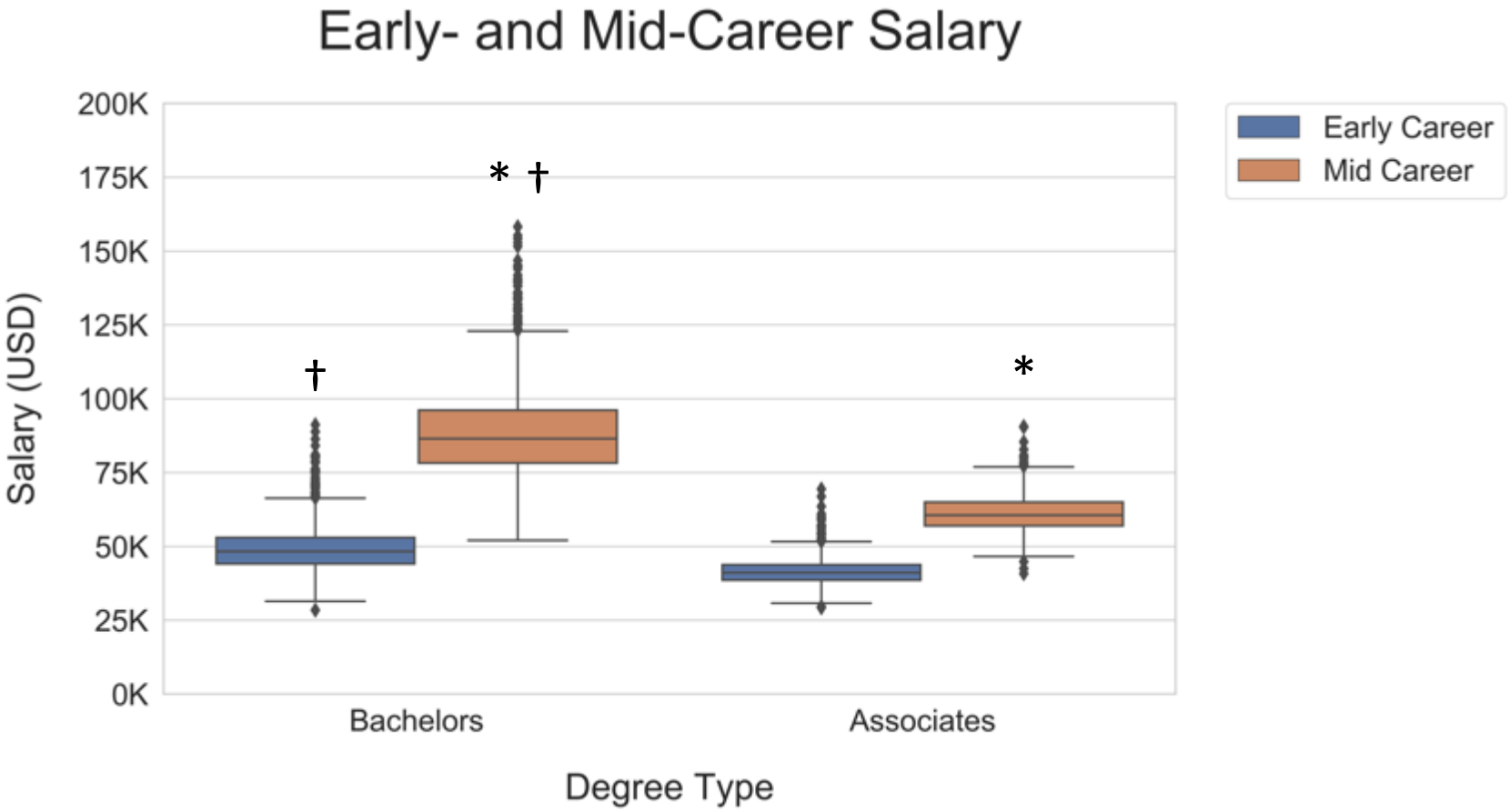


Figure. Early and mid-career salary by 4yr Bachelors degree and 2yr Associates degree. * indicates significant difference from early-career salary, per group. † indicates significant difference in salary between bachelors and Associates graduates.

ROI Distribution of In-State and Private Colleges, with and without Financial Aid Package

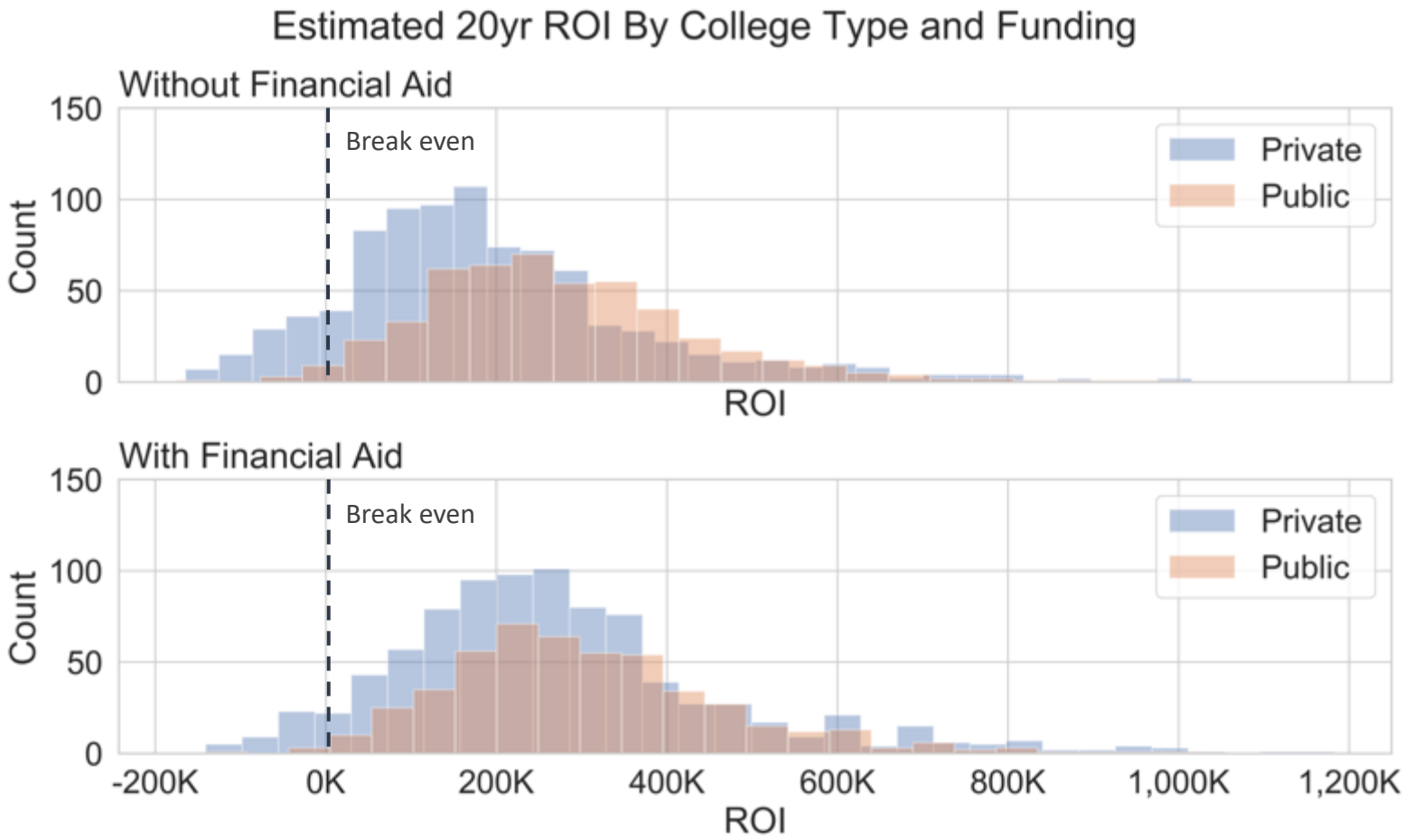


Figure. Count distribution of estimated 20yr ROI of Bachelor graduates. Out-of-state ROI and military service academies were excluded from analysis.

%Difference Between Private and In-State College ROI

Public ROI > Private ROI (p<0.05):

- AL, CA, KY, MI, OH, SC, TX, VA

Private ROI > Public ROI (p<0.05):

- PA

* 14 States with n<5 public or private schools excluded from statistical analysis

- Both <5: AK, AX, DE, HI, ID, MT, NV, UT, WY
- Public < 5: IA, NH, NM, ND, DC
- Private<5: NM

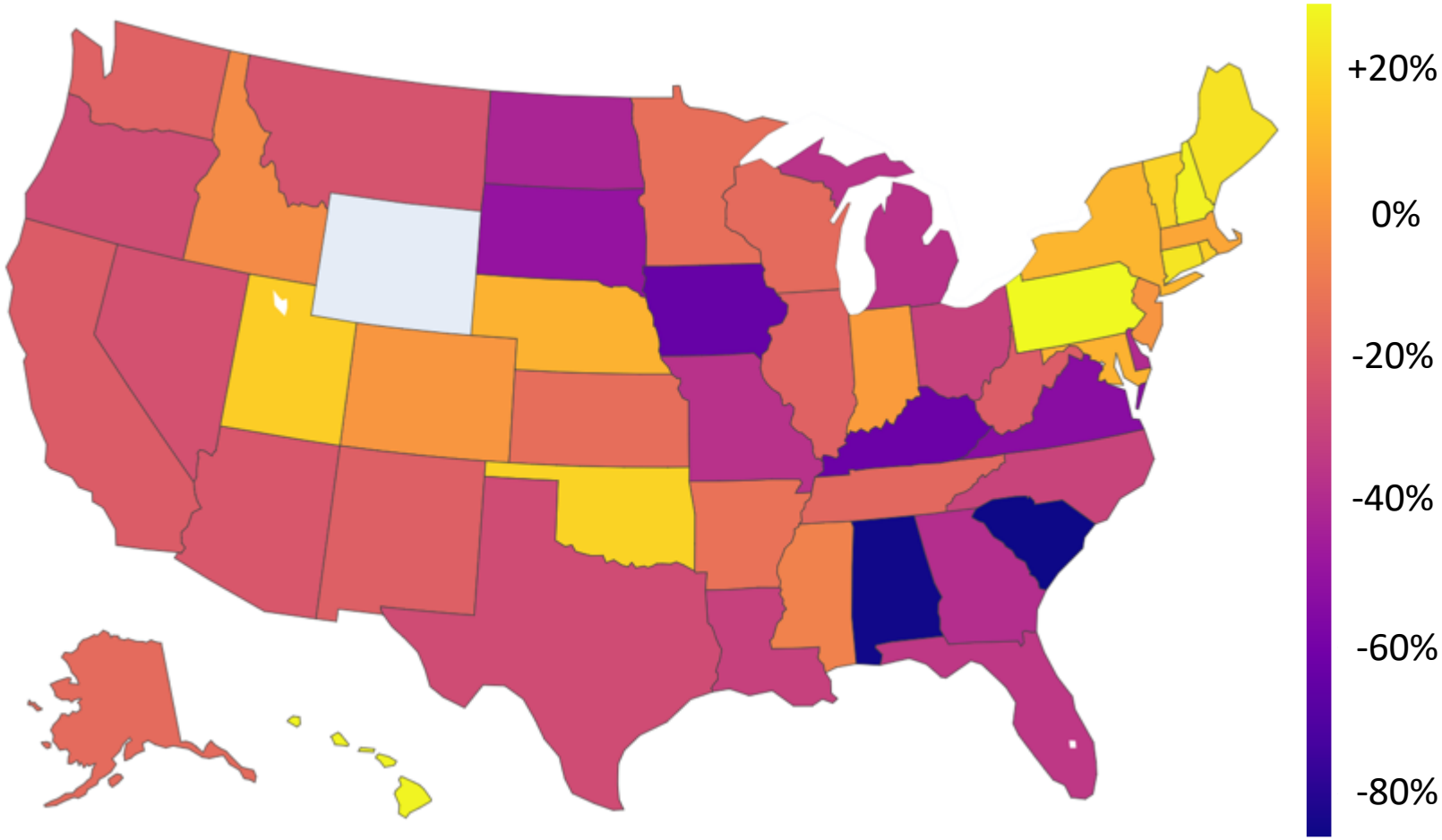


Figure. Percent difference in 20yr ROI between Private and In-State Public colleges. %diff > 0% indicates Private ROI > Public ROI. [Equation: (Private-Public)/(avg ROI)*100]

Salary Related to %STEM Majors at University

Early-Career

- Pearson correlation, r = 0.59 (p<0.05)

Mid-Career

- Pearson correlation, r = 0.60 (p<0.05)

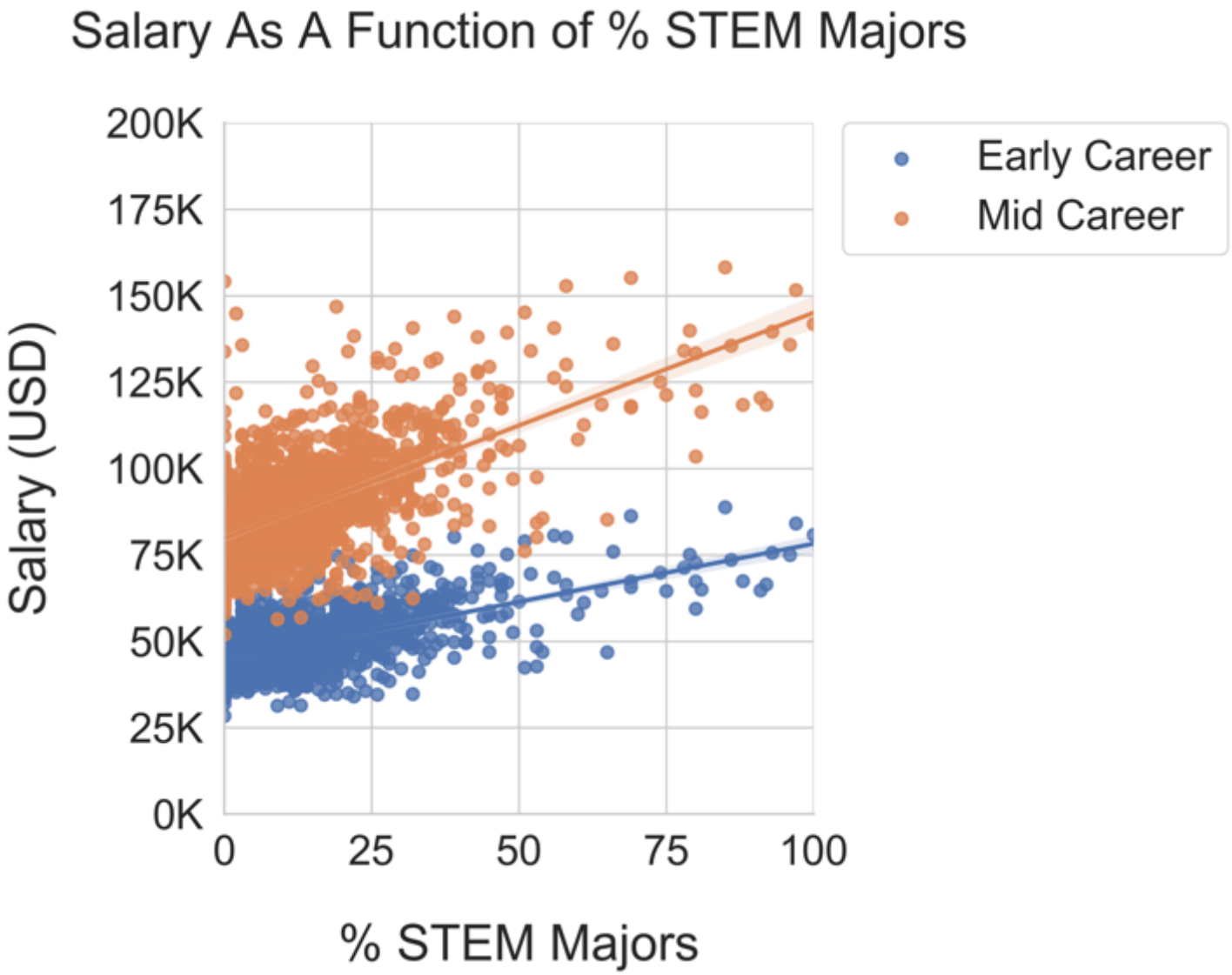
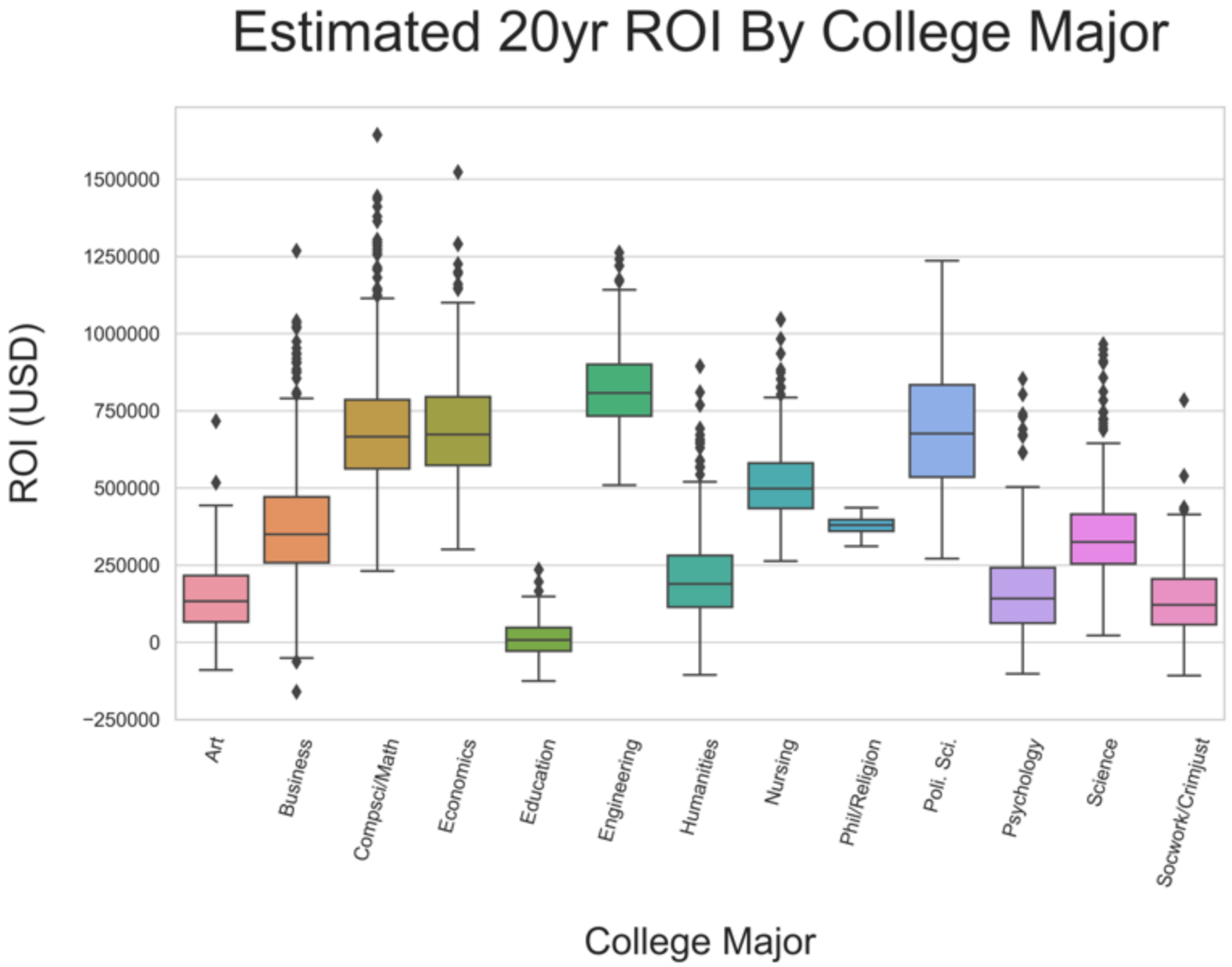


Figure. Early- and mid-career salary as a function of percent of student body graduating with STEM bachelor degrees.

ROI by College Major



ROI by College Major

Tukey Post Hoc Results

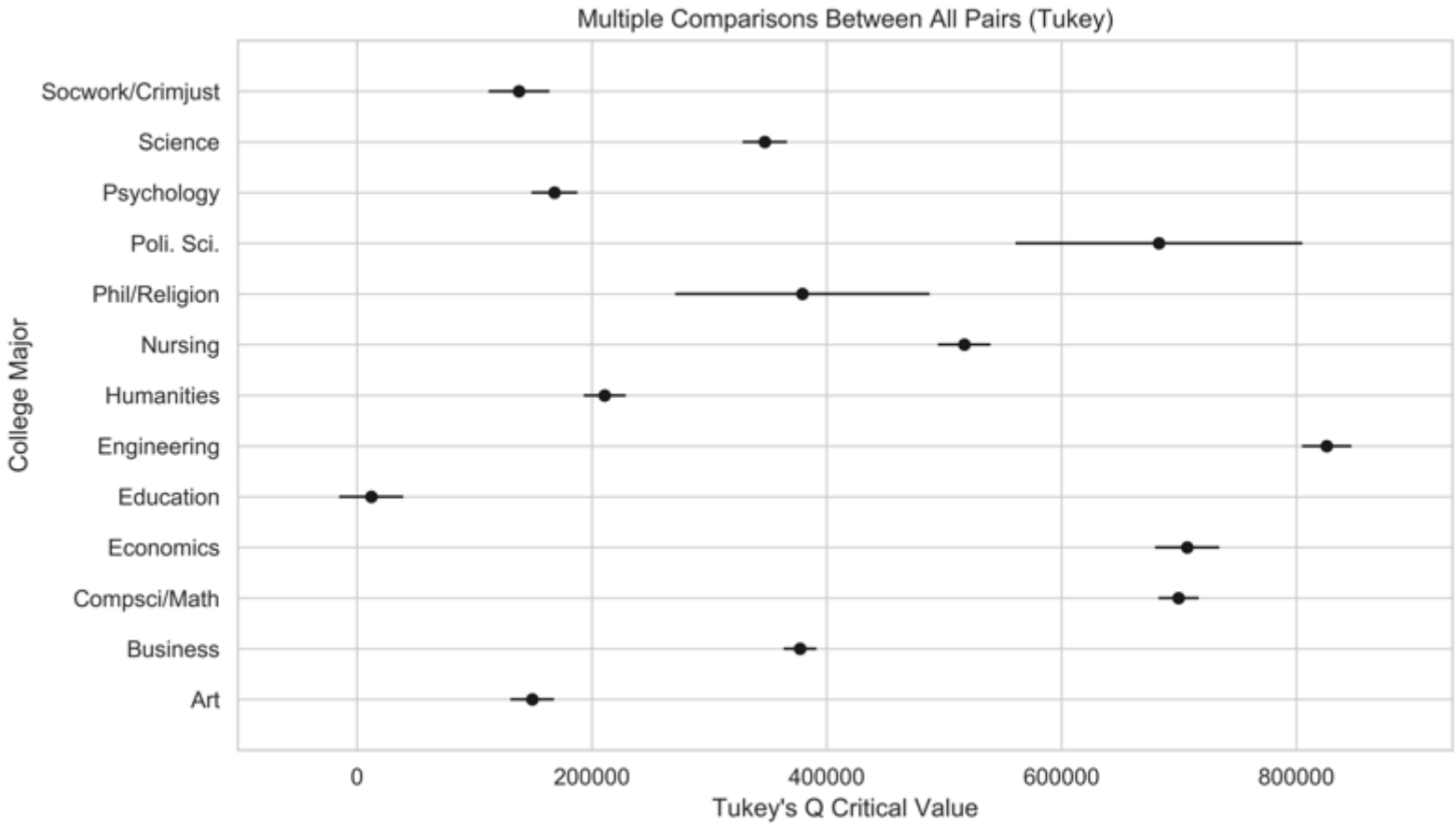


Figure. Statistical results of ANOVA Tukey's post hoc presented using Tukey's Q critical value to compute interval width. Significant relationships are defined as intervals that **do not** overlap.

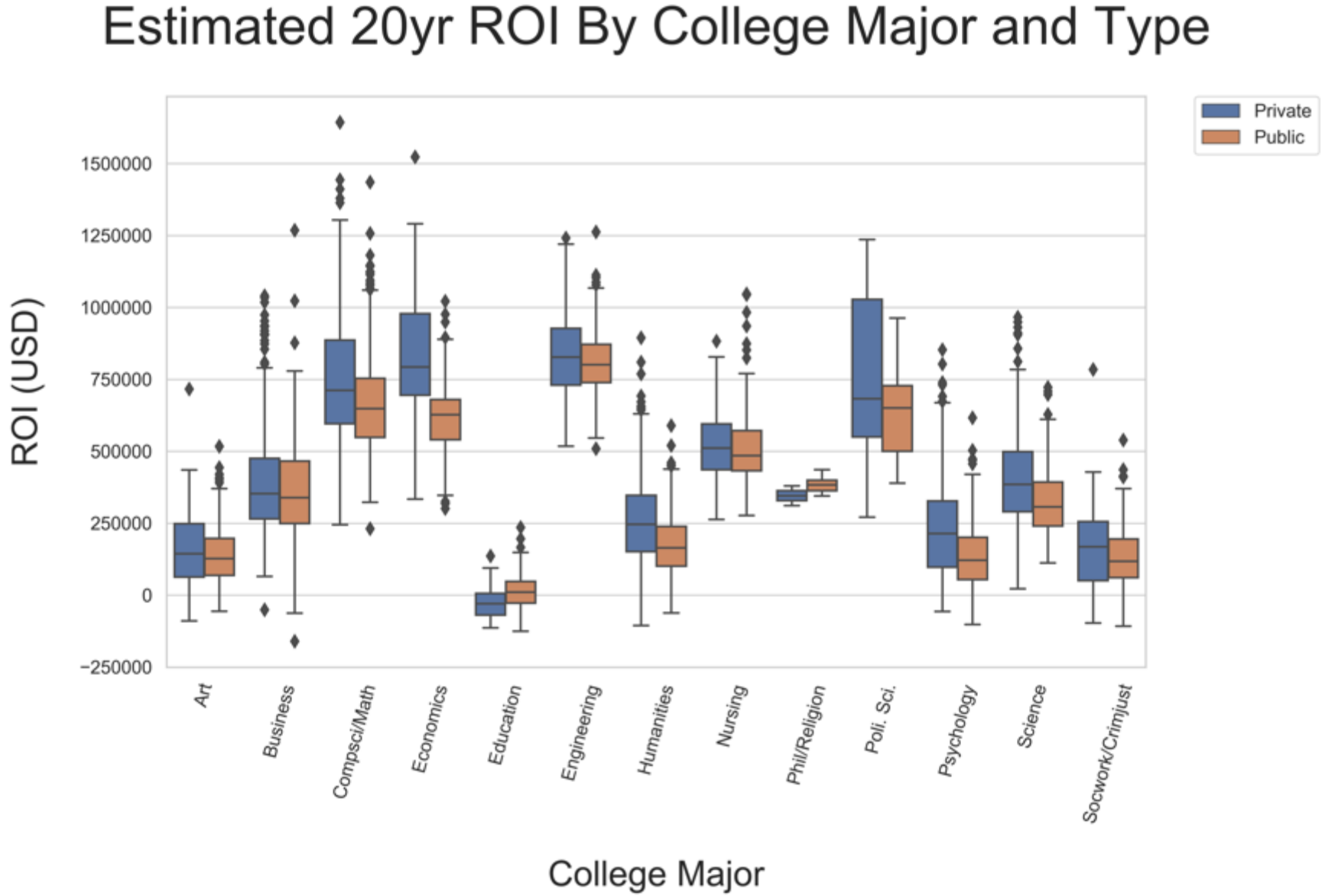


Figure. Estimated 20yr ROI per major for each private and public universities

Discussion

Graduates of 4yr colleges significantly out-earned their associate degree counterparts
- Earning 17% more within 0-5yrs; increasing to 36% in year 5-10
- Available data did not allow for 20yr ROI comparisons that would account for differences in education cost

Financial aid/grants subsidized as much as 11% of 4yr private colleges that would otherwise result in a negative ROI
- Only 2% of 4yr public universities returned a negative ROI without financial aid

Attending a higher-priced 4yr private university may only be financially beneficial in the north-east corridor
- Public universities in AL, CA, KY, MI, OH, SC, TX, VA significantly improved estimated ROI compared to private schools

Average early- and mid-career salaries of 4yr universities significantly correlated with %STEM of the graduating class
- Results emphasize the high-paying careers within the STEM field

Significant differences between estimated ROI exist between college majors
- Computer science, economics, and engineering majors had the highest estimated ROI of the majors

Differences between estimated ROI between private and public schools exist per major
- No significant difference were observed for art, nursing, political science, and social work/criminal justice

Limitations

While I stand by the analysis, I question its validity

Limitations

Estimated ROI takes into consideration:

- 20yr median pay for a bachelor's grad
- 24yr median pay for high school grad
- Total 4yr cost of the university

Estimation incorrectly assumes the cost of the university is the list price; doesn't account for interest

- Typical repayment period for borrowers with $20k-$40k = 20yrs [11]

Personal Case Study:

- Assuming I pay off my $100k in federal loan, at 8%, in 20yrs, the actual cost of my schooling is $200,600
- This would dramatically change the ROI of college compared to high school graduates, or public university

References

1. https://www.forbes.com/sites/zackfriedman/2019/02/25/student-loan-debt-statistics-2019/#46e28a31133f
2. https://www.usatoday.com/story/money/2019/05/18/cost-of-college-the-year-you-were-born/39479153/
3. https://www.epi.org/publication/charting-wage-stagnation/
4. https://ktla.com/2020/01/16/university-of-california-wants-5-straight-years-of-annual-tuition-increases/
5. https://www.usatoday.com/story/college/2017/06/09/private-college-tuition-is-rising-faster-than-inflation-again/37432483/
6. https://www.cbpp.org/research/state-budget-and-tax/state-higher-education-funding-cuts-have-pushed-costs-to-students
7. https://flatironschool.com/blog/coding-bootcamps-vs-college
8. https://www.payscale.com/college-salary-report/best-schools-by-state
9. https://www.payscale.com/college-roi/state
10. https://www.payscale.com/college-roi/major
11. https://studentaid.gov/manage-loans/repayment/plans/standard#monthly-payments-consolidation

Questions?

Contact Information

LinkedIn

- https://www.linkedin.com/in/jonharriseit

GitHub

- https://github.com/jah377

ResearchGate

- https://www.researchgate.net/profile/Jonathan_Harris23