

Explaining Startup Success by Industry

Hanna Ginzburg

December 2017

Abstract: This paper aims to isolate the propensity for startups within particular industries to obtain higher rates of funding, with success defined purely on financial terms. The methodology includes a multiple linear and logistic regression model to determine significant predictors of funding. The results suggest that startups catering to entertainment-related industries are associated with higher expected funding by 18.85% ($p=0.003$) and an increase in log odds of above-average funding of 0.936 ($p<0.001$) relative to other industries. Startups in the industry of sustainability are associated with highest levels of funding, by 26.3%, but at a statistically insignificant level ($p=0.122$). And, startups eliciting private equity funding, relative to seed funding, can expect higher funding by 55.8% ($p<0.001$) and an increase in log odds of above-average funding by 1.356 ($p=0.004$). The analysis does not discount technological innovation as a factor, but it demonstrates that the most important driving factor for funding may involve appeal to end users (whether or not the company's product or service is intended for user entertainment).

Introduction

Several articles have attempted to pinpoint the most important factors driving higher funding for startups. Tucker Marion draws on data from venture capital firm First Round Capital and asserts that geography actually doesn't matter much: "There's evidence that startup success is somewhat geographically diverse, not limited to Silicon Valley" (Marion). Another analysis by Bloomberg turned to Quid's CEO Mark Goodson, using artificial intelligence trademarked by his company to predict which 50 companies would be most successful. He began with a list of 50,000 that had received venture capital or venture debt in the past three years, using S&P Capital IQ and Crunchbase databases, and gave "preference to companies with... at least two rounds of funding; fewer days between rounds; fewer days since the last round; and founders who had worked together previously now bringing in a third outside partner" (Reddy). Over 50% of the companies chosen were based outside Silicon Valley, with 20% based outside American borders. The average business on the list had already raised three rounds, and the top sectors by capital raised were online security and fraud detection, artificial intelligence, autonomous driving, image recognition and mapping, augmented reality, smart sensors, drones, and digitization of education (Huet). Clearly, technology seems to be a vital factor determining the financial success of new startups; but, are there specific industry sectors that gain the most attention from investors? This paper aims to explore patterns of funding for new startups, as well as to identify the industries that receive above average funding, and conclude whether or not industry can be a statistically significant predictor of funding rates.

Methods

The dataset used in this analysis originates from a sample of 610 startups, whose information is listed on Crunchbase (up to date as of November 2017) mostly for the purpose of garnering attention from investors. All of the information is crowd-sourced, as people and companies enter data on themselves and others; most, however, includes links to press releases or news articles to confirm veracity. The variables included in the sample are log-transformed total funding amount, an indicator for above-average funding (average funding \$476.3 million), number of investors, number of lead investors, number of employees, number of funding rounds, funding timeline (duration from first to most recent funding round, in years), and indicators for headquarters location (grouped by predominant language spoken), indicators for funding status (seed, private equity, late stage venture, early stage venture, M&A, and IPO), and indicators for industry (the main variables of interest). Locations, or regions, were grouped by countries speaking European languages, countries speaking Asian languages, countries speaking Middle Eastern languages, as well as English-speaking countries and Spanish-speaking countries, with "other" (including countries like South Africa) used as a base. The mean of the linear dependent variable, log of total funding amount, is 19.27, with a maximum of 23.48 and a minimum of 18.35, and a standard deviation of 0.903.

I built two models, a multiple logistic regression and a multiple linear regression, in order to isolate the effect of each industry, along with the other predictors, on funding rates. I used a cutoff level of 5% to determine significance and R to conduct my analysis.

Results

Results of the multiple linear regression model appear in Table 1. The linear regression output indicates that companies producing goods or services related to the financial services sector and the entertainment sector are expected to have significantly higher funding rates. Specifically, a startup in financial services is expected to elicit higher funding by 11.38% ($p=0.03$), and a startup in the entertainment industries is expected to elicit higher funding by 18.85% ($p=0.003$), relative to "other" industries adjusting for other company characteristics.

The multiple logistic regression output yields fewer significant terms. The only industry group that shows up as a statistically significant predictor of startup funding in this case is entertainment; the difference in log odds of receiving above-average funding for startups in the field of entertainment versus “other” industries is 0.936 ($p < 0.001$), all else held constant. Startups related to the industry bins of business services, healthcare, and government are associated with negative log odds relative to “other,” though these three industries are not statistically significant.

Discussion

Interestingly, startups that create products and services catered to users’ entertainment appear to elicit highest funding. This “entertainment” industry bin includes any startup listed as germane to gaming, sports, video, media and entertainment, and/or music and audio. It seems that consumer demand is a positive driver for investor decisions, as games, music, video, and the like are always in high demand. Startups producing goods or services in the realm of government and military are weakly associated with lower funding rates -- possibly because this industry has not yet caught up to others on the basis of implementing technological innovation.

It is perhaps surprising that the most significant startup funding type for driving higher funding rates is private equity, with IPO trailing behind, relative to seed. A possible explanation may be that companies with initial public offerings have maximized their potential investments at a relatively low level, and are then forced to sell shares on the public market. Notably, funding timeline, number of founders, and location are not statistically significant, though the first two are weakly negative predictors.

Potential limitations of the data include the small sample size and potential selection bias. All Crunchbase startup data are entered by a self-selecting group, the startups included in the dataset all received significant amounts of funding to begin with, and funding totals in USD were calculated using exchange rates as of late November 2017. There may also be measurement error in the dependent and/or independent variables; if this is true for one or more predictor variables, we cannot negate the possibility of attenuation bias. Furthermore, there may be omitted variable biases on the basis of missing predictors, such as more detailed information about company founders and their business plans (charisma, wealth, level of education, attention to detail, up-front planning, and more).

Conclusion

It appears as though only one set of industries can be a statistically significant predictor of high funding rates; by the regression models, startups in entertainment-related industries are expected to elicit higher funding by 18.85% ($p = 0.003$) and an increase in log odds of above-average funding of 0.936 ($p < 0.001$) relative to other industries. Startups in financial services-related industries are also expected to elicit relatively weakly higher funding by 11.38% ($p = 0.03$). Startups making products or services related to the government, military, and/or business services are associated with lower expected total funding, though this is insignificant at a threshold of $p = 0.05$. Confirming Marion’s analysis, startup success does appear to be geographically diverse, rather than limited to English-speaking countries. For startups hoping to maximize total funding, these regression results show that an entertainment-related company idea pitched to a private equity source would be the optimal option.

Appendix

Table 1: Full regression output of the multiple linear regression model; dependent variable in this case is log of total funding.

Coefficients	Estimate	Std. Error	T Value	Pr(> t)
(Intercept)	18.04	0.23	77.48	< 2e-16 ***
Services	0.01	0.05	1.70	0.089
Business Services	-0.06	0.07	-0.85	0.397
Goods	0.05	0.08	0.56	0.576
Tech	0.03	0.03	1.05	0.294
Health Care	0.04	0.10	0.41	0.683
Sustainability	0.26	0.17	1.55	0.122
Education	0.06	0.02	0.38	0.708
Government	-0.39	0.41	-0.94	0.349
Science	0.00	0.05	0.02	0.983
Retail	0.02	0.07	0.21	0.833
Financial Services	0.01	0.05	2.14	0.033 *
Entertainment	0.19	0.06	3.04	0.003 **
European region	0.27	0.23	1.18	0.234
English-speaking region	0.21	0.20	1.03	0.305
Spanish-speaking region	0.03	0.75	0.04	0.968
Asian region	0.33	0.22	1.54	0.123
Middle Eastern region	0.42	0.46	0.92	0.359
Number of Investors	0.02	0.00	3.81	0.000 ***
Number of Lead Investors	0.08	0.02	4.51	7.96e-06 ***
Number of Employees	0.00	0.00	10.24	< 2e-16 ***
Number of Founders	-0.02	0.02	-0.69	0.488
Number of Funding Rounds	0.03	0.01	2.71	0.007 **
Funding Timeline	-0.00	0.00	-1.45	0.148
IPO	0.24	0.10	2.30	0.022 *
Early Venture	0.13	0.14	0.93	0.355
Late Venture	-0.04	0.08	-0.48	0.630
M&A	0.06	0.17	0.33	0.739
Private Equity	0.56	0.11	5.25	2.09e-07 ***
Signif. Codes:				
0 '***'	0.01 '**'	0.02 '*'	0.05 '.'	0.1 ' '
(581 degrees of freedom)	Multiple R-squared: 0.4255	Adjusted R-squared: 0.3978		

References

Huet, Ellen. "These Are the 50 Most Promising Startups You've Never Heard Of.

Bloomberg.com, Bloomberg, 6 Mar. 2017.

Marion, Tucker J. "4 Factors That Predict Startup Success, and One That Doesn't."

Harvard Business Review, 3 May 2016.

Reddy, Shashi. "A Computer Was Asked to Predict Which Start-Ups Would Be

Successful. The Results Were Astonishing." *World Economic Forum*, Quid, 21 July 2017.