The Global Accelerator Learning Initiative (GALI) exists to answer key questions about enterprise acceleration. As such, it is important to us to understand more about the ventures that are applying to accelerators. One question asked frequently is whether certain types of ventures are more likely to have received investment prior to applying to an accelerator program.

We challenged students from around the world to tackle this question for our 2017 Student Data Competition. Students used data from the Entrepreneurship Database Program at Emory University, which contains information from over 8,500 entrepreneurs that applied to nearly 100 different accelerator programs between 2013 and 2016. The winning submission was chosen due to its clear approach, quality analysis, and pertinent take-aways for the sector.

Several of the students' observations validated our current understanding of early-stage ventures. We were not surprised to see that ventures with established track-records and intellectual property are more likely to report outside financing. However, the students also uncovered some new trends in the data concerning the characteristics of founding teams. We found it particularly intriguing that the size of the founding team can predict the type of financing received, and that the age of founders and their prior work experience also plays a significant role in the likelihood of funding.

The students' work also emphasizes common concerns that team gender and the region in which the business operates significantly affects the likelihood of outside investment, showing a clear disadvantage in equity financing for teams led by women and for ventures based in sub-Saharan Africa.

As our dataset continues to grow, and interest in entrepreneurship expands within the academic community, we invite other students and researchers to join our efforts in expanding the knowledge base around early-stage entrepreneurship and acceleration around the world.

The GALI Research Team

Question

In our dataset of applicants to accelerator programs, what types of ventures (and entrepreneurs) are more or less likely to have received investment? Do different types of early-stage ventures attract investment from different sources?

Student Researchers

NISHANT GUPTA, MBA student at The Wharton School, University of Pennsylvania

PIYUSH KESHRI, MBA student at The Wharton School, University of Pennsylvania

HONGYE MAI, MBA student at The Wharton School, University of Pennsylvania

Summary

For startups, the recipe of being attractive to investors and raising financing (equity, debt, or philanthropic) is a healthy revenue stream, intellectual property, a social media presence, and a team of multiple founders. Regional differences also manifest while predicting the likelihood of financing across regions, with startups operating in Sub-Saharan Africa being less likely to raise any type of financing and those in Europe and Central Asia being more likely.

Other characteristics such as financial goals, team size, and founder age are also correlated with certain types of startup financing. An intriguing issue observed is that gender diversity in the founding team reduces the likelihood of equity financing significantly. Factors that are weak predictors of financing include the business model, impact areas, and profitability.

Approach / Methodology

RESPONSE VARIABLE

We observed that funding amount in the dataset is heavily skewed, with over 58% of ventures seeing no funding and ventures on the high end having received very large amounts. Moreover, regional differences make it challenging to compare amounts raised. As a result, we transformed investment received into a binary variable that indicates financing success. Overall, a firm is considered to have received an investment if it received any type of equity, debt, or philanthropy financing. We also constructed variables that indicate whether startups secured any of the three types of funding from the various listed sources (angel investors, companies, government, venture capital, etc.).

FEATURE ENGINEERING

We segmented the data features into three broad segments that we believe intuitively could impact a startup's likelihood of securing funding: founding team, business, and geography. Below are key transformations we performed on the raw data across these three categories.

1. FOUNDING TEAM

Number of founders: We devised three categories of startups:

- Single founder (1 founder)
- Small team (2-5 founders)
- Large team (more than 5 founders)

Age: Average age of founders in the startup

Women-led startup: Categorical variable capturing whether the first listed founder of a company is a woman or not

Female-male diversity: Ratio of the number of female founders to the total number of founders

Work experience: Cumulative years of founders' work experience prior to founding their current venture

Number of companies: Cumulative number of companies that founders have worked for (includes for-profit companies, non-profit companies, and other companies)

Kind-of-role: Categorical variable capturing the senior-most role that founders held prior to founding their current venture, in descending order of seniority are as follows:

- CEO
- Senior management
- Support staff
- Others

Education: Categorical variable capturing the highest education level of the founding team, grouped into five categories:

- PhD
- Master's degree
- Bachelor's degree
- High school or less
- Others (includes remaining sub-categories)

2. BUSINESS

Social motive: Indicates whether a company identifies as having the explicit intent of creating social or environmental impacts

Impact area: We grouped impact areas into six broad categories. A company is considered to have an impact in one of the following areas if it indicates impact in any listed sub-area.

- Economics (income; employment)
- Environment (energy; water; waste)
- Housing
- Health (disease; health improvement)
- Social (education; women & girls)
- Others

Business model: Includes the categories provided in raw data: production/ manufacturing, processing/packaging, distribution, wholesale/retail, services, financial services, and unsure

Intellectual property: Categorical variable indicating whether the company is invention-based or has patents, copyrights, or trademarks

Social media: Considers a company to have media presence if it has a website, Facebook, Twitter, or LinkedIn presence

Revenue: Separates revenue-generating companies from pre-revenue ones, using the maximum of reported revenue from the previous year and revenue since founding

Profit: Divided into three categories based on the startup's profit in the prior year:

- Low profit - negative to 5%
- Medium profit - 6-20%
- High profit - more than 20%

Financial goal: Indicates whether a company's financial goal is to cover cost or cover cost and earn some profit

3. GEOGRAPHY[1]

Venture operating region: Based on venture operating country

Founder region of birth: Based on most common country of birth among the founding team

Founders not local: Categorical variable indicating whether the venture operating country is different than the country of birth for the majority of the founding team[2]

1 We categorized countries into seven regions using the 2017 World Bank Country & Lending Groups – East Asia & Pacific, Europe & Central Asia, Latin America & Caribbean, Middle East & North Africa, North America, South Asia, and Sub-Saharan Africa.

2 We excluded following features from the analysis as they were highly correlated with the above-mentioned features: Venture headquarter country; founder residing country; founders' job country prior to founding the venture

REGRESSION ANALYSIS

Features from the above three segments were combined for the regression analysis, with "any financing" as the response variable. For each type of financing (equity, debt, philanthropy, and overall investments), we performed the following regression analyses:

- Ran a lasso regression on all the variables and selected the lambda value (shrinkage coefficient) that provided the best fit.
- Ran a logistic regression using the variables selected from the lasso regression.
- Further pruned the logistic regression by removing non-significant variables from step 2.

The logistic regression output estimates the probability of a startup receiving funding from the four response variables. The final regression models including all the statistically significant variables are as follows:

any_investment ~ no_of_founders + women_led + kind_of_role + edulevel + any_revenue + IP + info_has_socialmotives + social_media + impact_ health + model_prodmanuf + financial_goals + info_venture _region

any_equity ~ no_of_founders + founder_age + female_male_diversity + women_led + kind_of_role + edulevel + info_has_socialmotives + impact_ health + model_prodmanuf + model_procpack + model_finserv + any_revenue + IP + profit + financial_goals + venture _region + founders_not_local

any_debt ~ no_of_founders + founder_age + female_male_diversity + women_ led + total_com_started + kind_of_role + edulevel + impact_environment + impact_health + any_revenue + IP + profit + venture _region

any_philanthropy ~ no_of_founders + founder_age + total_com_started + edulevel + info_has_socialmotives + impact_econ + model_prodmanuf + model_services + any_revenue + IP + social_media + profit + financial_goals + venture_ region

Results

The above regressions include all the significant variables. We then selected the variables that were among the most significant and could bring insight to salient issues for founders and investors. The table below summarizes these key variables (ordered by statistical significance) that predict the likelihood of getting financed through different channels (equity, debt, and philanthropy).

Financing Likelihood

Any type (Equity / Debt / Philanthropy)	Equity	Debt	Philanthropy
↑ Revenue	↑ Revenue	↑ Revenue	↑ Higher education (Phd)
↑ Social media	↑ Financial goal (cover cost and earn profit)	↑ Medium Profit (6-20%)	↑ Social media
↑ Large team (more than 5 founders)	↑ Intellectual property	↑ Kind of role (CEO)	↑ Revenue
↑ Intellectual property	↑ Small team (2-5 founders)	↑ Average age of founders	↑ Large team (more than 5 founders)
↑ Operating region (Europe & Central Asia)	↓ First founder is a woman	↓ Country region (Middle East & North Africa)[3]	↑ Has social motive(s)
↓ Operating region (sub-Saharan Africa)	↓ Operating region (sub-Saharan Africa)	↓ First founder is a woman	↓ Average age of founders

3 Middle East & North Africa is the least-represented region in the data (44 ventures), so a larger sample may reflect different trends.

FURTHER DISCUSSION ON KEY VARIABLES

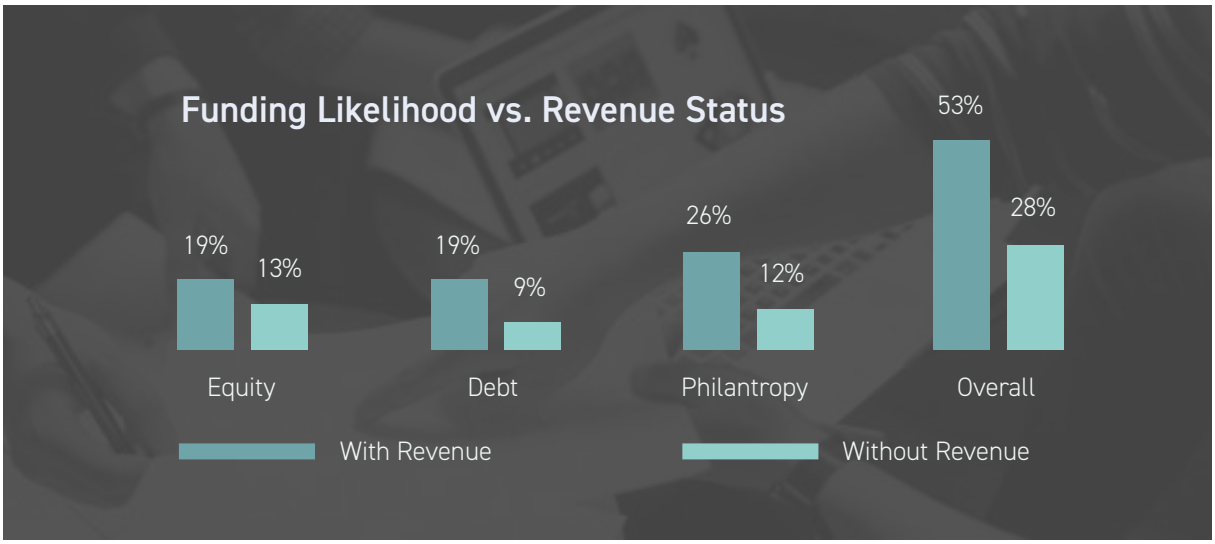
The following section further discusses the impact of several of the significant variables on the probability of funding. We discuss the regression results and use the existing dataset to check for any observable systematic relationships between these variables and the various types of funding.

BUSINESS

Revenue

Revenue is one of the most important variables that impacts funding likelihood across the board. A company that generates revenue has a significantly higher chance of receiving funding, and the effect of revenue is larger for debt and philanthropy financing. This is likely because revenue generation can serve as proof of financial viability where debt investors could be more confident of future repayment and philanthropy investors could be more confident in the business's sustainability. Equity investors, on the other hand, are typically focused on growth and can look past near-term delays in revenue generation.

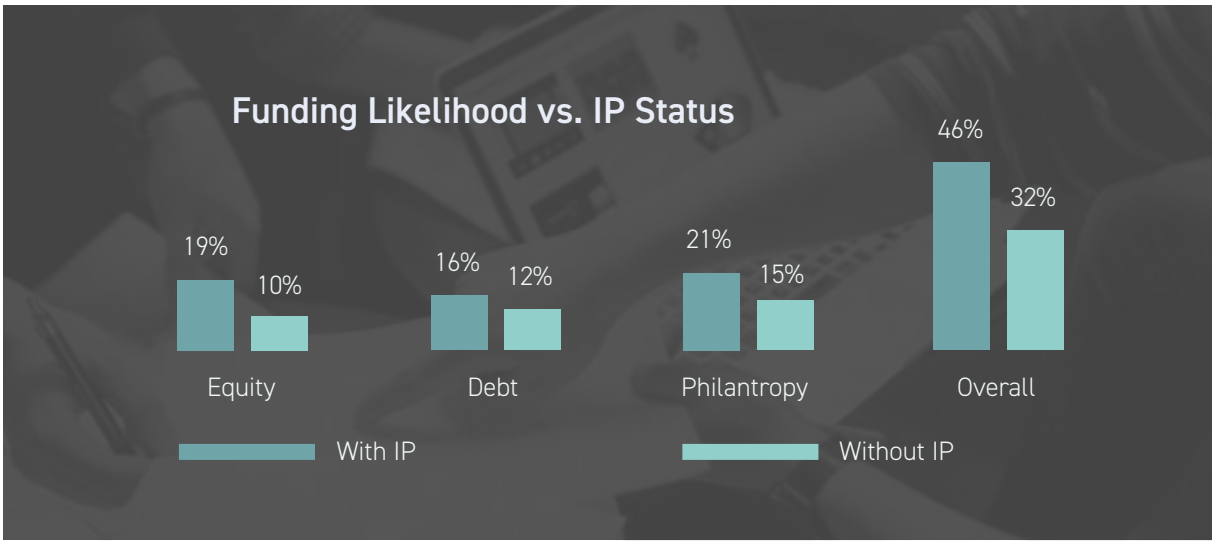
Regression analysis - Even after controlling for other variables concerning the founding team and region, revenue remains consistently one of the top three most significant predictors of each type of startup financing.



Intellectual Property

We defined intellectual property broadly as the startup having some fashion of invention, patent, copyright, or trademark. A company with intellectual property is more likely to receive funding. Intellectual property is particularly important for equity investors, as it typically serves as a key defense that sustains the startup's business model and is often the biggest value driver in acquisitions.

Regression analysis - After controlling for other variables, intellectual property is one of the top predictors for equity and overall financing. It is also statistically significant for philanthropy and debt funding.

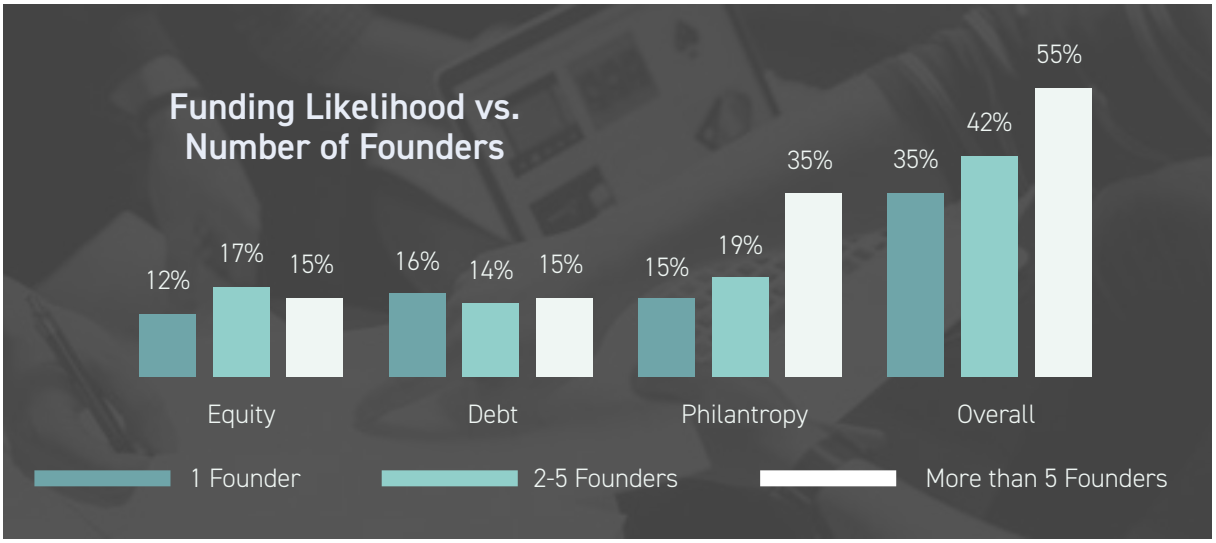


FOUNDING TEAM

Number of Founders

The size of the founding team is a major factor in determining the likelihood of funding. The higher the number of founders, the more likely a startup is to receive philanthropy and overall investment.

Regression analysis - However, a counter-intuitive statistically significant result is observed for equity funding, where the optimum number of founders lies between two and five. This could be explained by the fact that a larger number of founders tends to create a greater split in equity and potential disputes, which might not be favored in equity investing.

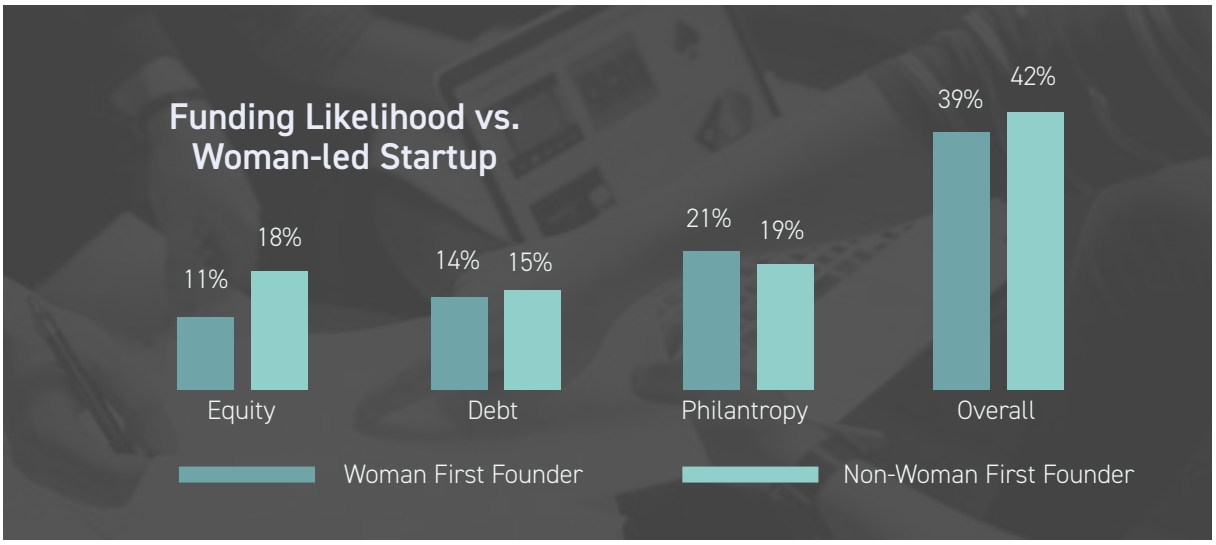


Women-led Startups

Whether a startup's first listed founder is a woman plays a statistically significant role in explaining whether a startup has received funding and the kind of funding it received.

Having a woman-led start up does not significantly change (increase or decrease) the chance of philanthropic funding; however, it does significantly decrease the likelihood of equity, debt and overall financing.

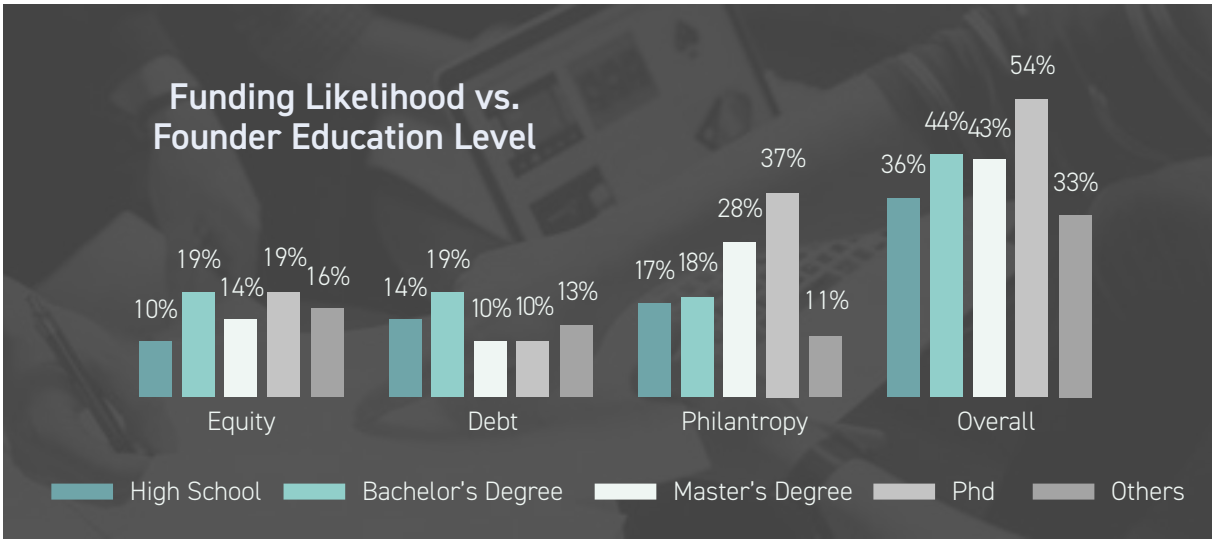
Regression analysis - The same result is observed with respect to the percentage of women on the founding team. As the percentage of women in the founding team increases, the probability of equity funding decreases (at a p level of 0.05), while the opposite holds true for debt funding.



Education Level

For overall and philanthropic financing, we observe that the probability of funding increases with higher education levels, with the highest being for PhDs. However, we do not see a strong correlation with equity and debt financing.

Regression analysis - We observe that having a founder with a PhD significantly increases the likelihood of receiving overall and philanthropic investment. Having at least one founder with a bachelor's degree increases the odds of debt and equity funding.



GEOGRAPHY

Entrepreneurs across geographies face different macroeconomic conditions to raise money for their ventures. We analyzed the funding scenario for startups operating across different regions.

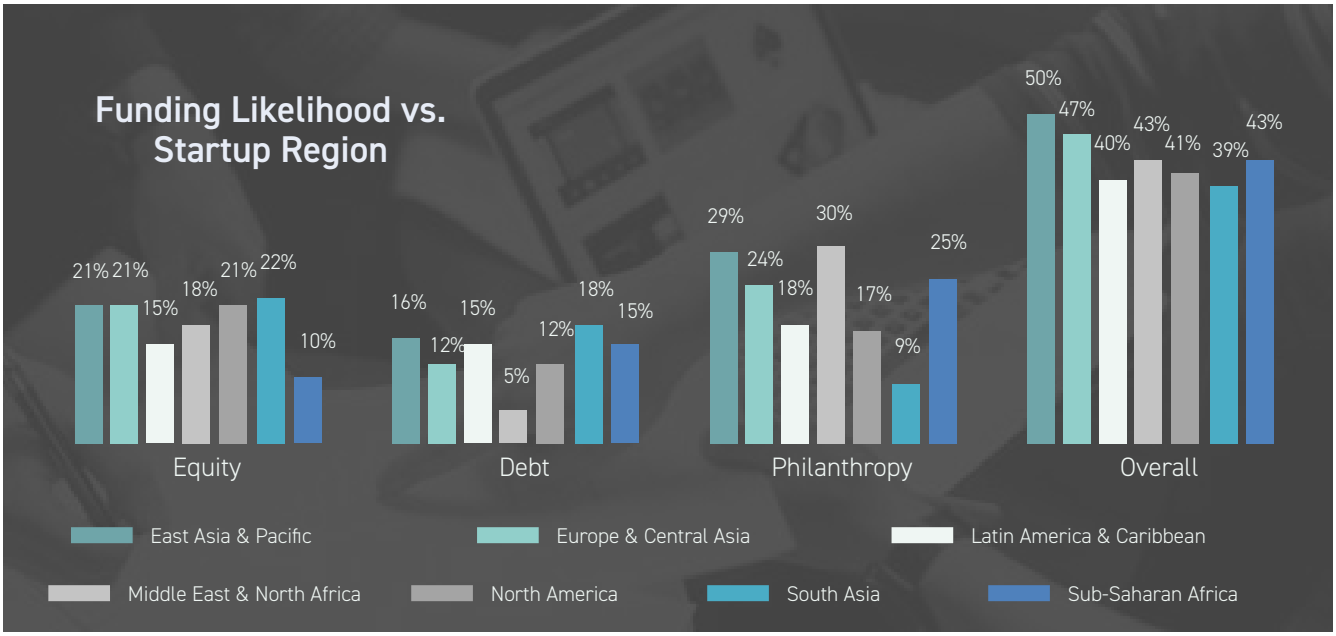
Overall financing - Roughly 40% -50% of startups were funded either by equity, debt, or philanthropy. East Asia & Pacific has the highest percentage of startups with financing (50%), while South Asia has the lowest (39%).

Equity financing - Sub-Saharan Africa has an abysmal success rate with equity financing (10%) compared to rest of the world (~20%).

Debt financing - The success rate of startups receiving debt financing across regions varies between 5% -18%. Regions do not vary distinctly for debt financing except Middle East & North Africa with the lowest debt financing likelihood.

Philanthropy - Philanthropy is much more prominent in the Middle East & North Africa and East Asia & Pacific (~30%). On the contrary, philanthropy is least common in South Asia (9%).

Regression Analysis – When controlling for other factors, any type of financing is limited in sub-Saharan Africa, especially equity. On the contrary, startups in Europe & Central Asia are significantly more likely to raise financing. Interestingly, the likelihood of receiving equity financing increases if the majority of founders are born in a different country than the location of the startup.



Conclusion

Analyzing data from startups applying to accelerators around the globe sheds light on funding preferences of various investors. While financing is not the only criteria to evaluate the success of a startup, it is crucial to its growth trajectory, long-term viability, and scalability. From our analysis, we observe several key trends that entrepreneurs and investors should keep in mind:

1. Business fundamentals are key to successful fundraising. A startup that generates revenue and has intellectual property to protect its competitive advantage will be favored by investors. Entrepreneurs should have an eye on developing intellectual property and be mindful of a path to monetization as they grow their startup.

2. The founding team is a strong indicator of startup's ability to navigate through uncertain times successfully. Solo founders face a difficult time receiving funding across the board. Having co-founders could not only improve the diversity of thought and camaraderie, it is also a preferred structure for investors. The appropriate education level in the founding team is also important. Having at least one highly educated co-founder can help improve technical sophistication and garner investor support.

3. Gender diversity in the founding team has an interesting effect on the funding opportunities available to a startup. While extensive studies have shown that diverse companies outperform their non-diverse counterparts, startups that are founded by women or have healthy gender diversity face challenges in raising money through equity and debt financing. This information is revealing about the potential bias women face in the startup world and should be brought to the attention of investors.

4. Entrepreneurs should be cognizant of the funding environment of their geographic region and areas of intended impact. While each startup is unique, none is insulated from its macro environment. Investors' risk appetite and growth expectations vary regionally as well as by impact area. It is important to identify the prevailing funding sources and direct financing efforts accordingly.

The Global Accelerator Learning Initiative (GALI), a collaboration between the Aspen Network of Development Entrepreneurs (ANDE) and Emory University, is designed to explore key questions about enterprise acceleration such as: Do acceleration programs contribute to revenue growth? Do they help companies attract investment? GALI builds on the Entrepreneurship Database Program at Emory University, which works with accelerator programs around the world to collect and analyze data describing the entrepreneurs that they attract and support.

Interested researchers are encouraged to download the data for their own analyses. Please visit www.galidata.org/about/ to learn more and to access the data.

The Global Accelerator Learning Initiative has been made possible by its co-creators and founding sponsors, including the U.S. Global Development Lab at the U.S. Agency for International Development, Omidyar Network, The Lemelson Foundation and the Argidius Foundation. Additional support for GALI has been provided by the Kauffman Foundation, Stichting DOEN, and Citibanamex Compromiso Social.