Causal Complexity Analysis for the Top European Startups

Kun-Huang Huarng[*]　Fang-Yi Lo[**]　Yehor Holikov[***]

Abstract

A startup is a nascent company that relies on innovative ideas and cutting-edge technologies. Startups are characterized by the inherent uncertainty of their success and the lack of sufficient financial resources during the creation or scaling stages of the project. This research examines the capitalization of startup companies and the investments that they attract within the funding rounds as a key measure of success. For statistical analysis, top European startups information from the open source database crunchbase.com was used. The findings of the fsQCA models constructed demonstrate that lead investors, technology adoption, active products, and website traffic ranking affect the capitalization of a startup. The study provides valuable insights for entrepreneurs and investors looking to improve their chances of success. By understanding the key success factors and how they interact with each other, startups can develop strategies to maximize their potential for success. Investors can use this knowledge to identify promising startups and make informed investment decisions.

Keywords: Startups, capitalization, technology adoption, investors, website traffic.

---

* Kun-Huang Huarng, Professor, Department of Creative Technologies and Product Design, National Taipei University of Business (E-mail: khhuarng@ntub.edu.tw)

** Fang-Yi Lo, Department of Management Science, National Yang Ming Chiao Tung University

*** Yehor Holikov, Master, Feng Chia Univeristy

Manuscript received: March 31, 2025; Accepted: August 14, 2025

臺北商業論叢　第3卷　第1期
民國 115 年 1 月　第 47-66 頁

歐洲新創企業之因果複雜性分析

黃焜煌[*]　　羅芳怡[**]　　柯易和[***]

摘　要

新創企業是一種依賴創新理念與頂尖技術的新興企業。新創企業的特徵包括其成功的不確定性，以及在創建或擴展階段缺乏充足的財務資源。本研究探討新創企業的募資情況及其在各輪融資中所吸引的投資，作為衡量成功的關鍵指標。研究所使用的統計資料來自開源資料庫，聚焦於歐洲頂尖的新創企業。因果複雜性模型分析結果顯示，主導投資者、技術採用程度、產品活躍情況以及網站流量排名，皆會影響新創企業的募資。本研究為創業者與投資者提供了寶貴的見解，有助於提升新創企業成功機率。透過探索關鍵成功因素及其交互作用，新創企業可制定合適之策略，以最大化其成功潛力；而投資者則能藉此識別具有潛力的新創企業，並做出更有根據的投資決策。

關鍵詞：新創企業、募資、技術採用、投資者、網站流量

[*] 黃焜煌，國立臺北商業大學創新科技與產品設計系教授
[**] 羅芳怡，國立陽明交通大學管理科學系教授
[***] 柯易和，逢甲大學國際貿易系碩士
收稿日期：114年3月31日；接受刊登日期：114年8月14日

I. Introduction

A startup is a company that relies on certain innovations, with a short history of operations and an expectation of future scaling. Forbes defines a startup as a business that has been just founded or started working, is innovative and disrupts traditional ideas about the development of its sector, has a real business plan, receives income that does not exceed the level of income of a startup, and is not acquired by companies, projects, or private foundations. In this research, we adopt Forbes' definition and consider a startup as an innovative business in the early stages of development with only a conceptual model but capable of rapid scaling.

For a considerable period of time, entrepreneurs and investors relied on their instincts and information gathered about similar businesses in the industry to predict the success of a venture. This approach proved difficult in the case of startups, given the innovative nature of their ideas. However, the emergence of Big Data has created new opportunities for predicting the success or failure of a startup, in addition to traditional entrepreneurial instincts and expert assessments. Nonetheless, as will be demonstrated, such analytical forecasts are often more effective in identifying areas where it is not advisable to invest than in providing reliable forecasts of success. Therefore, this study aims to identify and investigate the criteria for success of startups as innovative businesses, examine the factors and characteristics that affect their success at different stages of development, and explore the potential of predicting the success of a startup based on its characteristics and the environment. To achieve these objectives, this research contributed to the theoretical gap by three approaches: a review of the relevant literature and propose combination factors to the success of the startups, and an analytical approach of the fsQCA that involves statistical analysis, lastly, provide reliable opening data sources.

The study aims to investigate the startup phenomenon, identify success criteria, and explore the potential for predicting success of capitalizations based on various start-up characteristics. By understanding the factors that contribute to startup success, entrepreneurs and investors can make informed decisions and maximize their chances of achieving their goals. To guide the research process, the following research question is formulated: What are the key factors that influence the success of startups, and how can these factors be utilized to predict and enhance startup success?

This research question serves as a guiding framework for the subsequent sections, where an in-depth analysis of various startup characteristics and their impact on success will be conducted. By addressing this research question, the study aims to contribute valuable insights to the field of startup entrepreneurship and provide practical implications for entrepreneurs, investors, and policymakers.

II. Literature Review

A. The success of startups

The comprehensive analysis of the success of a startup is incomplete without a quantitative assessment of its value. In order to achieve an applicable assessment that is suitable for all types of startups, it is necessary to estimate the capitalization of startups.

Despite having a well-prepared business plan, a significant number of startups still fail within the first year of existence. According to CB Insights, five out of ten innovative companies go bankrupt in the first three months, and three more before the end of the first year of existence. The primary reasons for startup failure are a lack of demand, conflicts within the team, and insufficient funds. One current trend in startup development is the movement of technology companies to Silicon Valley due to its high success rate and proximity to centers of innovation and organizations.

The external environment is predominantly favorable for the growth of startups. Litau (2020) claimed that the two most influential factors for startups are a well-developed innovation infrastructure and high levels of private investment in the innovation sector, providing a conducive environment for startup development, even in the post-COVID-19 pandemic era. Felin et al. (2020) further examined the success factors for startups are: the ability to maximize the use of national and global innovation infrastructure, including access to venture capital; easy access to the results of science and technology parks worldwide; a high level of personal innovative developments, including a novel initial idea; a highly skilled team of specialists, with the potential to attract foreign talent; the ability to attract significant investments while maintaining attractiveness for investors. Among the various factors that contribute to startup success, two main drivers stand out: the demand for the business idea and the availability of adequate funding. These two drivers, in turn, depend on a broader range of factors (Thiel & Masters, 2015). It is important to note that both internal and external factors can influence either or both of these drivers.

Based on the previous literature, the success factors of a startup can be grouped into these primary components: First, easy access to financial resources, including substantial amounts of investors, and a well-developed workforce with high levels of team expertise.

Second, a high technological level and actively initiate the new products. Lastly, in the new era of digitalization, website traffic is a major issue affecting the visibility of this startup.

B. Investor team

According to Wasserman (2012), companies founded by teams are more likely to receive venture investment compared to companies founded by single entrepreneurs, with only 3,526 companies with one founder receiving investment to start their business. The number of founders will influence the degree of control, and the resources they can obtain then influence the value creation (Wasserman, 2017). It is important to note that the report's findings provide startups and potential investors with a set of common criteria for success that billionaire startups share. The lead investors will represent the reputation effect and attract more investors to join (Zhang et al., 2023). The first hypothesis proposes that startups with more lead investors are more likely to succeed than those with a larger or smaller founding team due to the team's composition and level of expertise.

Hypothesis 1: The number of lead investors in a startup is associated with its success in attracting investment.

C. Active Product

The lean start-up method iteratively designs a minimum viable product to learn about consumer preferences for new product ideas, has become popular among start-up entrepreneurs and intrapreneurs within large firms (Blank, 2013). As per the widely used Stage-Gate model (Ferrati & Muffatto, 2019) which represents the current innovative processes, the sequence includes four stages: the "pre-seed" stage, the "seed" stage, the startup stage, and the product's development resulting in either a positive outcome or the product leaving the market. At the "pre-seed" stage, the focus is on marketing the idea and establishing a relationship with the product's consumers which continues through the "seed" stage. The first step is to ensure that the product is in demand by consumers in a specific market, and then the product can be made scalable in other markets. The "seed" stage of a startup is primarily focused on developing a minimum viable product for the beginning of sales. At this stage, the emphasis of marketing is on the study and verification of the potential consumer of the product, as well as on the willingness of consumers to accept a new product. The innovative product should eventually be directed towards its consumers, and therefore the main task at this stage is to determine the content of the new product required, whether there are consumers in the market who will need the product, and who will be able to purchase it.

Modern lean startup techniques effectively develop the concept and content of a product being prepared for launch on the market (Alvarez, 2015). This approach enables obtaining estimates of the demand for the product prior to the start of sales. After analyzing the seed phase of a startup, it is important to consider the development of a product that leads to successful scaling. Scaling a business refers to its ability to increase profits through quantitative expansion, and it can be divided into horizontal, which involves adding new services or improving existing ones, and vertical, which involves growth in main indicators such as customers, products, services, and markets. The seed stage involves the formulation of a product idea and the development of a prototype. The startup stage requires the company to demonstrate a pilot version of the product and undergo product testing. In the early stage, the product is ready for market entry and undergoes demand testing. In the expansion stage, the product gains market acceptance, resulting in a surge in sales and demand. Finally, the late stage involves the transformation of the company into a large organization, indicative of a public company.

Alvarez (2015) developed a set of questions aimed at obtaining clear answers on who the consumers are, what their needs are, what problems they face, what influences their behavior, how they make purchasing decisions, and what non-existent products they are willing to pay for. The answers to these questions reflect the psychographic characteristics of potential clients, including their values on money, predictability, health, dependence on others' opinions, decision-making style, preference for leadership or obedience, innovativeness, incentive to buy, and concerns and problems to be solved. This approach allows for the identification of "early evangelists" (Sohl, 2011), or the pioneers who are the first to buy and adopt new products and are crucial for the success of innovation. These early adopters are the primary targets of sales efforts after the launch of the product (Ulwick, 2005).

The main conclusion of the research by Tripath & Oivo (2020) is the importance of launching a product or service at the right time, even if the idea seems promising. It is crucial to assess whether consumers are ready for innovation before launching. The second hypothesis posits that startups that are active with more products have a higher probability of success.

Hypothesis 2: The number of active products in a startup is associated with its success in attracting investment.

D. Technology

The research also did not negate the mandatory requirement for startups to have complex technological solutions and fundamental innovativeness.

Osterwalder et al. (2010) proposed an innovative business model, including the concept of lean development, which is applicable for preparing startups of innovative products. The main approach of the model emphasizes analyzing the alignment between the innovation and the consumers' wants and needs. Osterwalder and his colleagues argued that while product development can be painful, involving numerous revisions and alterations, developers may not have met their potential buyers and may not have considered their preferences or desires. Therefore, the model proposes analyzing the compatibility between the innovation and consumer need (Osterwalder et al., 2010). In conclusion, models for evaluating innovative startup ideas primarily focus on evaluating a specific startup idea, rather than identifying factors that increase the likelihood of success. While it is possible to evaluate an existing idea using these models, it is difficult to construct an idea that has a greater chance of success than other existing or potential ideas. Even hypothetically assuming the possibility of creating a model that generates more successful ideas, the different focus of startups in different lines of business remains a significant challenge.

The level of technical and scientific innovation ranges from pure system integration (where the added value is mainly in the business model) to high-tech companies where the main value is a new technology that is otherwise inaccessible (Tamaseb, 2021). Nonetheless, high-tech companies occupy a disproportionate share of the billionaire startup segment, leading venture capitalists to revise their risk assessment model and grant deep tech startups additional success rates. The third hypothesis posits that startups that adopt more trendy technology have a higher probability of success.

Hypothesis 3: The number of technologies currently adopted by a startup is associated with its success in attracting investment.

E. Website Traffic

The increasing availability of diverse big data and the development of mechanisms to obtain such data have opened up innovative approaches in the field of entrepreneurship. Many researchers have proven the positive relationship between web traffic and the value of a firm, but few have done so for startup companies and the different effect (Huang, 2020). The use of web analytics to improve online marketing dates when the first web analytics systems were developed (Keating, 2000). Plaza (2009) analyzed what will produce effective web traffic and found out return user visits are the mainstream sources of website visiting.

Predicting the probability of success of startups at an early stage of development has been a subject of considerable interest among researchers over the past decades. Many have attempted to determine how investors choose the best entrepreneurial projects to invest in. During the digital era, one of the main sources would be the internet. With the growing amount of information available about startups and venture capital funding, advanced data science techniques can now be applied to discover non-trivial, implicit, and potential star-ups to invest in.

Furthermore, a statistical analysis can be conducted to examine the relationship between the volume of attracted investments by the rating of global traffic or the monthly attendance of the company's web page. The former can be considered as an indicator of the level of attendance, and it can be used to study the impact of web traffic on the startup. There is a positive correlation between a startup's global traffic number and the amount of funding it receives. As a startup's website ranking position increases, the total amount of funding raised by the startup increases. The fourth hypothesis suggests a correlation between a startup's global traffic rank and its success, measured by the total amount of investments attracted.

Hypothesis 4: A startup's website global traffic rank is associated with its success in attracting investment.

After analyzing the literature on factors affecting the success of startups, several hypotheses can be formulated. The success of a startup depends on various factors, including the team composition of lead investors, market conditions of active products, technology adoption, and finally, the website traffic. While several methodologies for evaluating startups exist in the previous literature, the following section of this research contributes to identifying other factors that contribute to the commercial success of innovative startups by valid database and fsQCA statistic models.

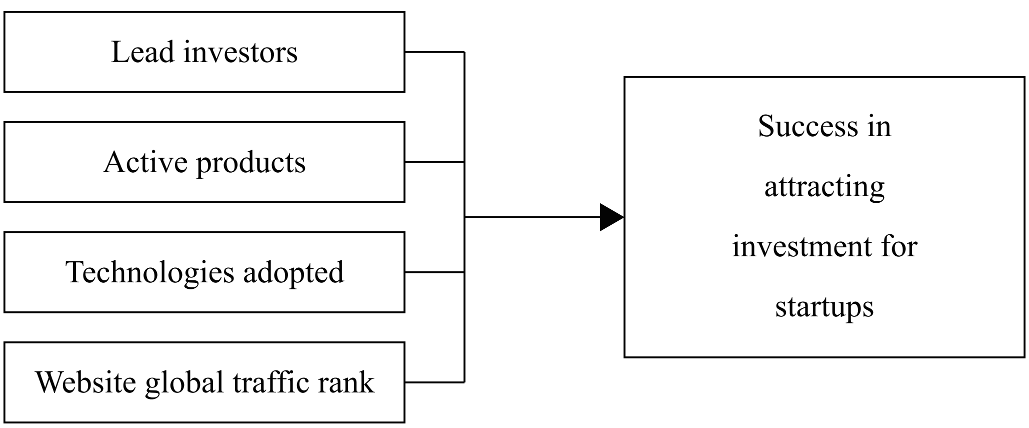

Figure 1. Research Framework

III. Data and research method

A. Data

The task of modeling a company at an early stage of development correctly is complex, and finding appropriate data is crucial for obtaining reliable results. In this regard, Crunchbase is an excellent source of information. It is an online platform that collects and provides business data on private and public companies globally. Originally created to track startups, the database includes descriptive data about companies, investors, funding rounds, and people involved in the entrepreneurial ecosystem. Sharing the data collection and validation strategies is a testament to the innovative approach and competitive advantage that Crunchbase has over other databases commonly used in the research field of entrepreneurship and economics. The quantity and quality of the collected data have led to the adoption of Crunchbase not only by practitioners, such as entrepreneurs, investors, and politicians, but also by academic researchers who employ a quantitative approach to studying entrepreneurship and innovation.

Crunchbase is a platform that provides information about businesses (Crunchbase, 2025). In this study, we use Crunchbase data for our analysis.

Crunchbase Inc. supports the platform and was founded in July 2007 by Michael Arrington in San Francisco, California. The project started as a subsidiary of parent company TechCrunch Inc., a well-known online publication established in 2005 by Archimedes Ventures (led by Michael Arrington and Keith Teer) focused on startups and the latest technology news. From 2007 to 2015, TechCrunch maintained control over the Crunchbase database, using it to track companies featured in articles. In September 2010, AOL acquired TechCrunch and Crunchbase as one of TechCrunch's portfolio companies for approximately $25 million.

In 2015, Verizon acquired AOL for $4.4 billion, and that same year Crunchbase spun off from AOL/Verizon to become an independent company (although AOL/Verizon, owner of TechCrunch, still retained a stake in the business). On September 22, 2015, in connection with the spin-off, Crunchbase announced a $6.5 million funding from Emergence Capital. On November 22, 2015, Salesforce Ventures, SV Angel, Felicis Ventures, Cowboy Ventures, and SVC raised a $2 million round of funding. On April 6, 2017, a round of funding of $18 million was announced with Mayfield as an investor. Felicis Ventures, Emergence, Cowboy Ventures, and AOL were among the investors. On October 31, 2019, the Series C funding round raised $30 million from OMERS Ventures, Mayfield Fund, Emergence, Verizon Ventures, Cowboy Ventures, and Felicis Ventures.

In recent years, Crunchbase has expanded its product offerings, introducing several new tools since 2016, including Crunchbase Pro, Crunchbase Enterprise, and Crunchbase for Applications in 2016, and Crunchbase Marketplace in 2018. The latest tool, Crunchbase Marketplace, has enabled the platform to integrate with third-party datasets to supplement its own information. Notable partner companies include Siftery, Apptopia, BuiltWith, SimilarWeb, IPqwery, Bombora, Owler, and Aberdeen. The demand for data provided by Crunchbase has also increased over time, as evidenced by search frequency data for the Crunchbase website worldwide.

In order to ensure the accuracy and reliability of data, it is crucial to consider the methodology used for collecting and validating the data in Crunchbase, which is partially a crowd-sourced database. Crunchbase employs four synergistic activities to collect, update, and validate its data on a daily basis. Firstly, through the Venture Program, investors provide monthly portfolio updates in exchange for a discount on Crunchbase API access, export to Excel, Crunchbase Pro, and Crunchbase Marketplace. This allows Crunchbase to access the most current data. Secondly, active members of the community can contribute information to the database, subject to registration, social verification, and moderation before being accepted and published. Thirdly, machine learning algorithms are utilized to validate data and identify inconsistencies. Lastly, the Crunchbase team of data scientists manually reviews the collected data and develops algorithms used internally to provide business insights and periodic reports.

B. Measurements

To measure the total funding of a startup, we use Total Funding Amount (Fund) as the outcome. This is also used in the previous work of Huang (2020). In Crunchbase, Total Funding Amount represents the cumulative sum of money a firm has raised across all its funding rounds.

The antecedents include:

Number of Lead Investors (LInvestor) – Lead investors will have a reputation effect to the new startups (Zhang et al., 2023). The total number of lead investors backing a startup.

Total Products Active (Product) – The number of active products a firm is currently using or offering. Number of active products can represent the startups built new ideas based on the customer’s need (Sudhir et al., 2025).

Active Tech Count (Tech) – Technology adoption by a startup will influence its financial performance (Huang, 2020). The total number of technologies a firm is currently using, as detected by BuiltWith.

Global Traffic Rank (Traffic) – The number of website visitors can represent the website traffic (Plaza, 2009; Huang, 2020). A web traffic ranking that indicates a website's popularity relative to all other sites on the internet. A higher Global Traffic Rank means a site has less traffic, making it a reverse indicator for our analysis.

To adjust for this reverse relationship, we calculated the Reverse Global Traffic Rank (RTraffic) as follows:

RTraffic = the maximum Global Traffic Rank − Global Traffic Rank

To analyze the factors influencing startup funding, we propose the following research model:

Fund = f(LInvestor,Product,Tech,RTraffic)

where higher values of Product, Tech, and RTraffic are expected to positively influence funding attraction.

C. FsQCA

FsQCA uses Boolean algebra to identify causal relationships (or combinations of antecedents) that contribute to the outcome of interest (Boswell & Brown, 1999; Ragin, 1987, 2000, 2009). FsQCA is widely used in management and business research because it offers significant advantages over regression-based methods (Woodside, 2013). One key advantage is its focus on complex and asymmetric relationships between outcomes and antecedents, whereas regression-based methods primarily compute the net effects of factors within a model (Pappas & Woodside, 2021; Rihoux & Ragin, 2009; Rihoux et al., 2013).

Ragin (1987) stated that fsQCA is a substitute method to identify a cause-and-effect process. Following Huarng (2015), fsQCA uses set theory to evaluate the correlation between antecedents (independent variables) and outcome (dependent variable) and take into consideration the antecedent or the combination of antecedents as a sufficient condition for the outcome. Our paper adopts fsQCA to achieve the optimal combination between TMT diversity attributes and TMT conventional compositions that can lead to better firm performance.

The first step in fsQCA is to calibrate the data into values between 1.0 and 0.0. Calibration ensures that all variables are transformed onto a common scale, regardless of their original measurement units. Pappas & Woodside (2021) suggest using the 95th, 50th, and 5th percentiles of each antecedent and outcome as calibration thresholds: Values above the 95th percentile are calibrated to 1.0 (full membership). Values below the 5th percentile are calibrated to 0.0 (full non-membership). Values at the 50th percentile are calibrated to 0.5 (crossover point).

The next step is truth table analysis, which enhances the validity and rigor of fsQCA results (Pappas & Woodside, 2021). Truth table analysis applies Boolean algebra for qualitative comparisons. Initially, the truth table includes all logically possible combinations of antecedents, where each antecedent is either present (1) or absent (0). When there are N antecedents, there can be up to 2^N possible combinations.

However, not all combinations need to be considered in the final analysis. The next step is to remove low-frequency combinations that do not provide sufficient empirical evidence. This study follows prior research by setting the consistency threshold at 0.80.

FsQCA generates three types of solutions:

Complex Solution – Includes all possible combinations of antecedents.

Parsimonious Solution – A simplified version of the complex solution, removing redundant conditions.

Intermediate Solution – Incorporates counterfactual analysis to balance complexity and parsimony (Liu et al., 2017; Ragin, 2009). The intermediate solution is often preferred in fsQCA studies because it integrates theoretically and empirically meaningful conditions derived from both complex and parsimonious solutions.

D. Handling missing data

Our dataset contains a significant amount of missing data, and the issue worsens as we move from top-ranking firms to lower-ranking ones. Table 1 illustrates this trend, showing that the number of missing values increases for firms ranked 101–200 compared to the top 100 firms, with an even more severe decline in data availability beyond this range. Given these challenges, this study focuses on the top 100 firms as the target sample.

Table 1. The amount of missing data

	Fund	LInvestor	Product	Tech	RTraffic
1-100	2	17	44	3	25
101-200	2	46	62	3	50

Despite this restriction, missing data remain a concern within the top 100 firms. To address this issue, we leverage the fsQCA calibration process. In fsQCA, data are transformed into values between 0.0 and 1.0, where:

1.0 represents full membership (absolute presence).

0.0 represents full non-membership (absolute absence).

0.5 represents a crossover point, indicating a neutral position between presence and absence.

Given this framework, we use 0.5 as an imputed value for missing data for two key reasons: Cases with a 0.5 value in any antecedent are ignored in combination generation – fsQCA excludes such cases when forming causal configurations, ensuring that missing data do not create artificial or biased combinations. FsQCA still considers 0.5 in consistency calculations – While ignored in combination formation, cases with 0.5 contribute to the overall consistency score, preserving the integrity of the analysis.

The consistency of a condition X as a subset of an outcome Y is computed using the following formula:

Consistency = $\sum \min (Xi,Yi) / \sum Xi$

where: Xi is the calibrated value of case i in the condition; Yi is the calibrated value of case i in the outcome; min is the minimum function.

By applying this approach, we ensure that missing data are handled systematically without distorting the results of the fsQCA analysis.

IV. Empirical analysis

A. Empirical results

The first step in the analysis is to calibrate all data. The 95th, 50th, and 5th percentile thresholds for the outcome and each antecedent are listed in Table 2. For example, the calibration thresholds for Fund are 5,203,200,000.00, 98,000,000.00, and 6,000,000.00, respectively.

For LInvestor, the thresholds are 5.90, 2.00, and 1.00. As a result, when the raw data for LInvestor equals 2.00, it is calibrated to 0.5. However, a large number of observations were assigned a 0.5 value, which may impact the analysis. To address this, we adopt a pragmatic approach by slightly adjusting the threshold from 2.00 to 1.90, reducing the proportion of cases assigned a 0.5 value. Table 3 presents the original data alongside their calibrated values.

Table 2. Calibration thresholds for the outcome and all the antecedents

	Fund	LInvestor	Product	Tech	RTraffic
95%	5,203,200,000	5.90	53.75	85.00	9,801,974.40
50%	98,000,000	2.00	19.00	36.00	8,761,204.00
5%	6,000,000	1.00	9.75	4.80	1,784,620.40

Table 3. Data and calibrated results

Name	Fund	Linvestor	Product	Tech	RTraffic	c_Fund	c_LInvestor	c_Product
1	2.75E+09	5	18	56	9.15E+06	0.83	0.91	0.42
2	1.32E+09	5	14	49	9.66E+06	0.67	0.91	0.16
3	1.10E+08	4	24	68	9.75E+06	0.5	0.83	0.61
4	2.46E+08	2	29	85	6.75E+06	0.52	0.52	0.7
5	1.20E+10	3	43	126	9.84E+06	1	0.7	0.89
6	1.59E+08	7		53	9.51E+06	0.51	0.98	
7	2.11E+09	5	10	45	9.57E+06	0.77	0.91	0.05
8	6.10E+09	5	72	95	9.85E+06	0.97	0.91	0.99
9	1.37E+08	2	34	69	9.67E+06	0.51	0.52	0.78
10	3.98E+08	2		38	9.47E+06	0.54	0.52	

For handling missing data, we impute the missing data using 0.5, following the above methodology. Table 4 presents the truth table, where we set the consistency threshold at 0.80 for identifying meaningful configurations. The fsQCA results are summarized in Table 5, showing the following solution:

c_LInvestor*c_Product*c_Tech*c_RTraffic

This indicates that high values across all antecedents (lead investors, active products, technology adoption, and reverse traffic rank) contribute to higher total funding for firms.

Table 4. Truth table with imputation 0.5

c_LInvestor	c_Product	c_Tech	c_RTraffic	number	c_Fund	raw consist.	PRI consist.	SYM consist
1	1	1	1	9	1	0.802982	0.473227	0.487805
1	0	1	1	6	0	0.790217	0.381873	0.383011
1	0	0	1	1	0	0.783758	0.326347	0.326347
1	0	0	0	4	0	0.773205	0.310254	0.315718
1	0	1	0	3	0	0.771133	0.278986	0.28361
1	1	1	0	2	0	0.754787	0.188139	0.188139
1	1	0	1	1	0	0.736182	0.211244	0.211244
1	1	0	0	1	0	0.722747	0.187404	0.189441
0	1	1	1	4	0	0.677555	0.121982	0.126984
0	0	1	0	2	0	0.659651	0.030702	0.03177
0	0	0	0	1	0	0.639906	0.064417	0.065217
0	1	0	1	1	0	0.626622	0.062726	0.062726

Table 5. fsQCA results for 0.5 imputation

Model: c_Fund = f(c_LInvestor, c_Product, c_Tech, c_RTraffic)			
--- COMPLEX SOLUTION ---			
frequency cutoff: 1			
consistency cutoff: 0.802982			
	raw coverage	unique coverage	consistency
c_LInvestor*c_Product*c_Tech*c_RTraffic	0.45681	0.45681	0.802982
solution coverage: 0.45681			
solution consistency: 0.802982			
--- PARSIMONIOUS SOLUTION ---			
frequency cutoff: 1			
consistency cutoff: 0.802982			
	raw coverage	unique coverage	consistency
c_LInvestor*c_Product*c_Tech*c_RTraffic	0.45681	0.45681	0.802982
solution coverage: 0.45681			
solution consistency: 0.802982			
--- INTERMEDIATE SOLUTION ---			
frequency cutoff: 1			
consistency cutoff: 0.802982			
Assumptions:			
	raw coverage	unique coverage	consistency
c_LInvestor*c_Product*c_Tech*c_RTraffic	0.45681	0.45681	0.802982
solution coverage: 0.45681			
solution consistency: 0.802982			

B. Discussions

Handling missing data effectively is a critical issue in both academic research and practical applications. This study adopts the “neither nor” value (0.5) in fsQCA as an imputation method. This approach offers a balanced solution: On one hand, fsQCA excludes cases with 0.5 when generating combinations of antecedents, ensuring that missing values do not distort causal patterns. On the other hand, fsQCA includes 0.5 values in consistency calculations, allowing these cases to contribute to the overall analysis. This balance justifies the decision and highlights a practical way to handle missing data in fsQCA-based studies.

Despite this contribution, handling missing data remains an ongoing research challenge. Future studies can further refine imputation techniques, offering more robust solutions for researchers and practitioners. From a business perspective, one of the key objectives for startups is to attract more funding. This study demonstrates that startups with: more lead investors, a higher number of active products, greater use of technologies, and higher web traffic, are more likely to secure higher funding. In supporting hypotheses 1 through 4.

Table 6. Hypothesis testing results

Hypothesis	Content	Results
H1	The number of **lead investors** in a startup is associated with its success in attracting investment.	Supported
H2	The number of **active products** in a startup is associated with its success in attracting investment.	Supported
H3	The number of **technologies** currently **adopted** by a startup is associated with its success in attracting investment.	Supported
H4	A startup's **website global traffic rank** is associated with its success in attracting investment.	Supported

This finding enhances existing literature by providing empirical evidence on the factors influencing fundraising success. It also offers practical insights for entrepreneurs and investors looking to optimize their strategies for securing venture capital.

V. Conclusion

A. Concluding Remarks

This research analyzes the factors of the startup as primary factors affecting its capitalization. Overall, the findings of this study suggest that a range of factors, including lead investor team size, active products, technology adoption, and website traffic ranking, can influence startup success. The study provides valuable insights for entrepreneurs and investors looking to improve their chances of success. By understanding the key success factors and how they interact with each other, startups can develop strategies to maximize their potential for success. Investors can use this knowledge to identify promising startups and make informed investment decisions.

B. Limitations and future research suggestions

We propose several areas for future research, which can also help address some of the limitations of the current study given the scope of our empirical analysis. First, the founder's branding effect may also influence the startups capitalization. Due to the data availability limitation, and not easy to use the objective proxy to measure the branding effect of the founders. Suggest the future study to conduct a subjective questionnaire to capture this issue. Second, not only the founder's roles, the roles of top management team may also influence the startups founding sources or follow-up IPO process. Although most of the samples in this study still have not gone so far and the current database did not provide the top management team's information. Future study may move further to explore the top management team's role over startups founding rounds to the next IPO stage.

C. Theoretical Implications

The current study's findings align with the previous literature on startup success factors. For instance, previous studies have suggested that the number of founders, technology, and products are crucial factors that affect a startup's success. The present study validated these factors. Furthermore, this study identified a new factor, a startup's global traffic rank, as a predictor of success.

The theoretical contribution of this study is twofold. First, the study offers empirical evidence that supports the validity of the previously suggested factors, demonstrating that they are not just anecdotal but are valid predictors of startup success. Second, the study contributes a new factor, a startup's global traffic rank, as an essential factor that determines startup success. These findings can guide investors, founders, and policymakers in their decision-making and investments.

D. Practical Implications

The approaches used in this study can be useful for both startup founders and venture investors. According to the literature review identifies main factors that significantly influence the capitalization of startups: the demand for and scalability of the core business idea and the success of financing rounds - from seeking business angels in the early stages of the company's operation to entering the IPO and subsequent public offerings of shares.

The study's results can guide investors to assess the potential success of a startup based on the lead investors, technology, and the products. Moreover, investors can use a startup's global traffic rank to understand the startup's online visibility and its potential for attracting investments. This can help investors in making informed decisions when investing in startups, which can reduce their risks of failure and maximize their returns.

The findings also offer guidance to founders, helping them identify the essential factors that can influence their startup's success. By identifying the industry trends, founders can strategically position their startup and tailor their services to cater to the current market demand. Moreover, founders can focus on increasing their online visibility by improving their website traffic, which can lead to more investments.

Finally, policymakers can use the study's findings to design policies that promote startup success. For example, policymakers can create programs that offer training and funding for startups with two to three founders, as they are more likely to succeed. Furthermore, policymakers can promote innovation and entrepreneurship in trending industries to support the growth of startups in these sectors.

References

Alvarez, C. (2015). *How to create a product that will be bought: Lean customer development method*. Alpina Publisher.

Blank, S. (2013). Why the lean start-up changes everything. *Harvard Business Review, 91*(5), 63–72.

Boswell, T., & Brown, C. (1999). The scope of general theory: Methods for linking inductive and deductive comparative history. *Sociological Methods & Research, 28*(2), 154–185.

Crunchbase. (2025). *Crunchbase*. https://www.crunchbase.com

Crunchbase. (2025, July). In *Wikipedia*. https://en.wikipedia.org/wiki/Crunchbase

Felin, T., Gambardella, A., Stern, S., & Zenger, T. (2020). Lean start-up and the business model: Experimentation revisited. *Long Range Planning*, 53(4), 101889.

Ferrati, F., & Muffatto, M. (2019). A systematic literature review of the assessment criteria applied by equity investors. In *Proceedings of the 14th European Conference on Innovation and Entrepreneurship* (pp. 304–312).

Huang, W. (2020). *It is all about attention: An analysis of web traffic effect on fintech startups* (Doctoral dissertation, Waseda University).

Huarng, K.-H. (2015, July). Re-examining the consistency in fsQCA. In *Annual Conference of the Global Innovation and Knowledge Academy* (pp. 102–109). Springer International Publishing.

Keating, E. K. (2000). Discussion of the eyeballs have it: searching for the value in internet stocks. *Journal of Accounting Research*, *38*, 163–169.

Litau, E. Y. (2020). Scoring method as applied to innovation project evaluation for start-up support. *Entrepreneurship and Sustainability Issues, 7*(4), 2978–2990.

Liu, Y., Mezei, J., Kostakos, V., & Li, H. (2017). Applying configurational analysis to IS behavioural research: A methodological alternative for modelling combinatorial complexities. *Information Systems Journal, 27*(1), 59–89.

Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, game changers, and challengers*. John Wiley & Sons.

Pappas, I. O., & Woodside, A. G. (2021). Fuzzy-set qualitative comparative analysis (fsQCA): Guidelines for research practice in information systems and marketing. *International Journal of Information Management*, 58, 102310.

Plaza, B. (2009). Monitoring web traffic source effectiveness with Google Analytics: An experiment with time series. *Aslib Proceedings, 61*(5), 474–482.

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. University of California Press.

Ragin, C. C. (2000). *Fuzzy-set social science*. University of Chicago Press.

Ragin, C. C. (2009). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago Press.

Rihoux, B., & Ragin, C. C. (2009). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Vol. 51). Sage.

Rihoux, B., Álamo-Concha, P., Bol, D., Marx, A., & Rezsöhazy, I. (2013). From niche to mainstream method? A comprehensive mapping of QCA applications in journal articles from 1984 to 2011. *Political Research Quarterly, 66*(1), 175–184.

Sohl, J. (2011). *The angel investor market in 2010: A market on the rebound*. Center for Venture Research.

Sudhir, K., Yoo, O. S., & Zhou, Z. (2025). Entrepreneurs and investors: Funding-induced distortions in lean start-up product experiments and innovation. *Marketing Science*, 1–21.

Tamaseb, A. (2021). *Super founders: What data reveals about billion-Dollar startups*. PublicAffairs.

Thiel, P., & Masters, B. (2015). *From zero to one: How to create a start-up that will change the future* (2nd ed.). Alpina Publisher.

Tripathi, N., & Oivo, M. (2020). The roles of incubators, accelerators, co-working spaces, mentors, and events in the start-up development process. In *Fundamentals of software start-ups: Essential engineering and business aspects* (pp. 147–159).

Ulwick, A. (2005). *What customers want: Using outcome-driven innovation to create breakthrough products and services*. McGraw Hill.

Wasserman, N. (2012). *The founder's dilemmas: Anticipating and avoiding the pitfalls that can sink a start-up*. Princeton University Press.

Wasserman, N. (2017). The throne vs. the kingdom: Founder control and value creation in start-ups. *Strategic Management Journal, 38*(2), 255–277.

Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research, 66*(4), 463–472.

Zhang, Y., Scholes, L., Fu, K., Hughes, M., & Tang, F. (2023). Equity crowdfunding syndicates and fundraising performance: The effect of human capital and lead investor reputation. *Journal of Small Business and Enterprise Development, 30*(4), 645–666.