WISEflow
Europe/Oslo(CEST)
30 Jun 2022

Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	16-01-2022 09:00	**Termin:**	202210
Sluttdato:	01-07-2022 12:00	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202210||10936||IN00||W||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Christian Anker og Paul David Maxime Valseth Helle

Informasjon fra deltaker

Tittel *:	GRA 19703 Master Thesis
Navn på veileder *:	Sheryl Winston Smith

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	187
Andre medlemmer i gruppen:	

Frontpage

Examination code and name:

GRA 1970 Master Thesis

Thesis

July 1st, 2022

Supervisor: Sheryl Winston Smith

BI Norwegian Business School

Master of Science in Business

Major in Strategy

Paul David Helle & Christian Anker

Acknowledgments

Our thesis work has provided challenges and rewards. We are grateful for our supervisor, Sheryl Winston Smith, who provided us with much-needed guidance. She deserves special thanks. Further, our families and partners for their unwavering belief and support in our work. Lastly, we are thankful for BI Norwegian Business School and our fellow students, who provided us with a supportive environment. We have learned a lot, and most importantly, we have grown. If we have glimpsed further, it is because of the shoulders we had to stand upon.

Oslo, July 1st, 2022

Paul David & Christian

Executive summary

Our thesis focuses on the strategy phenomenon of accelerator programs and how they emerged to meet global demand for increased innovation and value creation. Our thesis explores the role accelerator programs play in catalyzing innovation in the energy sector. We seek to understand which network factors are vital in influencing the accelerator program network and subsequently enhancing innovation. Innovation studies continue to be central in solving humanity's many challenges. With an aim to do network development research, we base our analysis on the theoretical perspectives of Network Theory and the Ecological view. We perform a comprehensive literature review to enhance our understanding of the phenomenon and verify our research question through gap-spotting, then derive hypotheses and shape our descriptive quantitative research design. The data we use is a novel hand-built dataset utilizing secondary data on three distinct energy accelerator programs retrieved from the public database Pitchbook. The completed data set was analyzed through an association study and triangulation. We uncovered systematic relationships in our data, and our findings confirmed most of our hypotheses. Comparing our findings to the existing literature aided us in discovering similarities and gaps. Notably, we found that network factors derived from characteristics within the network, such as participant group features, hold a great influence on the accelerator program performance. Whereas startups and investors have been found to geographical cluster, our study finds that the use of accelerator programs increases startup-investor distance. We suggest implications for research and practice. Based on our limitations, we propose improvements to research design and areas worthy of future research. We conclude that our study has added value to the strategy field. Lastly, our findings are used to offer recommendations to energy accelerator programs.

Table of content

1.0 Introduction

Science is the process of producing generalized understanding through systematic observation (Straits & Singleton, 2018). Our thesis intends to expand on what is known, push boundaries, and contribute to the collective understanding of corporate strategic decision-making. We intend to study the phenomenon of *network effects on accelerator programs*. The network itself is the phenomenon under study, and we view accelerator programs as innovation networks central to performant innovation ecosystems. We argue our phenomenon is of great importance to our collective understanding of the mechanics of modern innovation practices.

Our paper flows logically and is structured based on our emerging research. Having decided on our research topic, we choose appropriate theoretical perspectives as lenses through which to view our phenomenon. An extensive review of the literature follows, assisting us in identifying uncovered areas and gaps on which to formulate our research question and appropriate hypotheses. Our research design and execution of it follow. We analyze our collected data material with an appropriate method and present our findings. They are discussed in light of the existing knowledge, and we derive implications for theory and practice. Based on limitations and experience, we propose areas for future research. The paper concludes with recommendations.

We chose our two theoretical foundations, *Network Theory,* and the *Ecological View,* based on their complementarity and ability to explain our phenomena. Network theory describes networks' and their participants' interactions, whereas the Ecological view depicts their behavior as a larger ecosystem in an evolutionary state. In combination, we get two dimensions with enhanced explanatory power through which we interpret our phenomenon under study. To be precise, our main lens is Network Theory, with the Ecological View to support it.

Our research question grew out of a longer investigative process. Gap spotting in the literature led us to ask several preliminary questions and initiate preliminary pilot studies to unveil if the direction had enough available data to be pursued. Our focus narrowed to accelerator program networks. Through our literature review, we discovered no articles that used the same angle as our question, which led us to believe our direction might uncover novel findings and provide valuable insights into the knowledge stream on our phenomenon. We ask: *Which network factors yield increased performance for an energy industry accelerator program?* Based on our research question, we formulated hypotheses we thought would describe and explain the network factors' effect on accelerator programs in general and energy accelerator programs specifically. By testing these hypotheses, we intend to uncover insights that hopefully can assist accelerator programs in optimizing their performance in the future.

To represent our research question, we located three similar yet distinct, energy accelerator programs to use as our units of analysis. They are perceived globally to be the top-performing programs. Their distinction is found in their organization method, where each utilizes a separate approach to achieve the same goal of successful energy-startup acceleration. We have an Industry partner accelerator ('IPA'), a Dedicated corporate accelerator ('DCA'), and a Dedicated industry accelerator ('DIA'). We aim to analyze them through a triangulation approach to uncover which distinct network factors lead to superior performance and use them to control for each other.

We use a quantitative approach allowing us to define a measurement that captures a larger sample size without lowering reliability or validity. Based on a lack of pre-made data sets covering our units of analysis, we use a public database to self-assemble a novel data set with what we argue are strong descriptive variables for analysis to help us test our hypotheses. The data processing is performed with the software tools Tableau and Python, and our analysis is as well.

Our research method is called *network development research* on the interorganizational level (Carpenter et al., 2012). That implies that we study causes that affect a network, in which the focal actors are organizations that combine to form a network. We try to recognize patterns and determinants of network formation and change from the network and non-network constructs working as *predictors* (causes). Therefore, particular network features become the *predicted effects* (consequences). To analyze the causes and consequences, we perform an *association study* (Altman & Krzywinski, 2015; Borgatti et al., 2009; Breiger, 2004; Carrington et al., 2005; Crossley, 2010).

Our study arrives at a good time. The world economy has never been as vibrant, and the rapid exchange of information in a globalized world has led to an unprecedented amount of innovation. We argue that understanding what leads and drives innovation is vital in our society to keep up the pace. Our thesis is important since it attempts to explain parts of how modern innovation happens and what strategic decisions could be made by organizations to improve upon the process to increase performance. Accelerator programs are a relatively recent creation whose success came to dominate new business creation in a short amount of time. They are now one of society's main vehicles for drastic innovations. Despite their success, the accelerator program model is a work in progress. Our approach offers to apply network theory and ecosystem theory to aid accelerator programs in configuring and tuning their innovative networks.

We have chosen to study energy accelerators for a reason. The ability to innovate has led humanity into a golden age of prosperity. The standard of living and quality of life by all metrics on a global scale has never been as high throughout history. The unfettered consumption of our environment's resources has, however, left us with two very prominent challenges. The first is our complete dependence on the hydrocarbon resources that continue to support our growth and way of life. However, such resources are finite, meaning that it is imperative that we find replacements before they run out. The second problem is that these resources' extraction, consumption, and waste have severe negative externalities on the environment. There has been a collective failure to tend to the global collective good, and, to the extent we do not change, humanity and life on the planet may face an existential threat. As the buildup of these externalities becomes apparent, and our reserves of these vital resources are diminishing, it puts increasing pressure on all actors to pursue sustainable and renewable solutions. The energy industry is at the center of these issues and is, therefore, our unit of analysis. It toils to achieve a green shift in operations with the intent to propel mankind away from fossil dependencies into more sustainable energy sources. Optimization of energy accelerator programs may decrease the time the world needs to get there by increasing the pace of innovation.

To comply with APA-7, we notify about the possibility of self-plagiarisms. Our paper may or may not incorporate parts, sentences, or wording used in our previously delivered papers in an exact, re-written, or paraphrased form. The thesis work has been a long time in the making, and we built key concepts and ideas upon each other as we went along. Based on APA-7 rules, we have not mentioned our previous works in the reference list.

2. Literature review

The purpose of our literature review is to present a critical review of what is known today about our phenomenon in light of our chosen theoretical perspectives (Straits & Singleton, 2018). To increase the probability of our thesis uncovering novel findings, we search the existing literature for gaps. We keep in mind that our approach was investigative, which means we were ready to change our premises should we discover sounder logic and reason during the search. After the literature review, we critically discuss the main findings to identify knowledge gaps. Then we derive hypotheses.

Scope and methodology

Through the 20th century, the business strategy field has worked to explain firm performance and survivability. A few significant schools have emerged, such as the industrial-organizational perspective (Porter, 1979); the activity-based perspective (Porter, 1996); the resource-based view (Barney, 1991); resource dependency theory (Pfeffer & Salancik, 1978), transaction cost theory (Williamson, 1991); organizational learning (Levinthal & March, 1993); the knowledge-based view (Grant, 1996); the ecological approach (Bruderl & Schussler, 1990); network theory (W. Powell, 1990; W. W. Powell et al., 1996; Uzzi, 1997, 2020); and institutional theory (Scott, 2001).

To build our network development research, we found Network Theory and the Ecological View to provide a good foundation in helping us explain how our phenomenon behaves. We see these theories as naturally intertwined. They are, in our view, good overarching metatheories with an ability to partly explain the other theoretical schools. Network theory describes networks' and their participants' interactions, whereas the Ecological view depicts their behavior as a larger ecosystem in an evolutionary state. They appear to be very complementary in this way. Most studies we have come over have previously used lenses whose scope is limited to the firm or industry level. As such, we hope that using these two more holistic scopes may provide new insights looking at the network and ecosystem levels. In our focus on innovation networks, such as accelerator programs, we look to the external environment organizations are part of and where they will source their new capabilities and resources. This is where network participants meet the accelerator programs. Network theory explains a firm's advantage as dependent on which network one is a part of and one's position within it to get the access one needs to perform and survive. The ecological view states that firm advantage comes from possibilities accessible in their encompassing ecosystem. To successfully balance the trade-offs to become long-term innovators, organizations must know how to leverage explorative and exploitative innovations. Such trade-offs are better made by having an understanding of the causes and consequences of network factors on the network.

We made our search scope *innovation, networks,* and *accelerators* in connection with *network theory* and *the ecological view*. To make our search more precise, we defined innovation, networks, and accelerators. Innovation as the "commercialization of an invention" (Ahuja, 2000; Schumpeter, 1939). The definition excludes inventions that fail to become available to use for the greater good while at the same time recognizing that invention is a necessary step in the innovation process. Networks as a "distinct form of coordinating economic activity" (W. Powell, 1990). Excluding other types of non-economic networks. Accelerators as "programs that accelerate certain factors in companies" (Mallaby, 2022). We were able to exclude other types of accelerators, most prominently physical particle accelerators. These definitions helped us refine our scope again.

Then we used Web of Science and Google Scholar to search for articles and books from highly cited authors in highly regarded journals and publishers to make sure our findings would be part of the knowledge stream within its field of strategy. Next, we data mined the reference lists of more recent articles to extend our search area to find quality articles on the field's border that might hold applications for our research. To be included, works had to be relevant to business strategy theory and come from esteemed strategy journals, such as Academy of Management Journal, Academy of Management Review, Administrative Science Quarterly, Journal of Management, Strategic Management Journal, Organization Science, and Research Policy. With the application of search strings, we found articles and books with the keywords in their abstract, title, and more. Our strings looked similar to this one: “Innovation" OR "Radical" OR "disruptive" NOT "Incremental" OR "Network theory" OR "Ecological theory" OR "Ecological view" OR "Ecosystem" AND "energy". Our search yielded results from the last hundred years.

We applied search restrictions to the Web of Science for themes, journals, and times cited. Still, our search yielded a large number of articles. We had to filter it down by scanning titles, abstracts, and cross-citations. Next, we read abstracts and grouped the articles before further synthesizing within each category, where we only chose those with the best fit to our phenomenon and perspectives.

To view patterns of change in preferred external innovation methods, we performed a Google Ngram Viewer search for word popularity in books and articles from 1900 to 2019. We discovered how *corporate venture capital* (‘CVC’) recently appeared in the limelight, gaining traction from around 2000. *Venture capital* (‘VC’) has been more permanently around since the 1960s and still dwarfs CVC. To gain perspective, we search for *merger* discovered as a buzzword growing through the 1900s, with a steep decline from 2000 onward. Lastly, we searched for the word *acquisition*. The most used by far, though in decline from the 1990s. The only search word with a notable increase from 2000-2019 is CVC. Lastly, we search for *accelerator* and found it to explode during the 1960s and the 1980s before plummeting until around 2010, when it gains new popularity. Since accelerator also refers to a particle accelerator and other such technological inventions, we believe this has distorted the search. Referring back to our previous definition, we find that accelerator programs became popular after 2010. The other words we searched did not have numerous interpretations. To further tap into the vast amounts of data on the internet, we performed the same search on the Reddit database, one of the largest chatter communities, with a GitHub program (Kashcha, n.d.) in hopes of generating a network of associated word clouds. Apparently, strategic capital allocation choices and innovation methods are not talked much about on Reddit, and the search delivered nothing of value.

Network Theory

Network theory spans interaction between individuals, organizational units, and firms in a web of relations. Its scope is both macro, industry, and firm-level. Here the "groups" or whom you interact within both the horizontal and vertical dimensions are both a source of potential competitive advantage and operational constraints. The network of a firm may result in exclusion from others. It is both a constraint and a social construct. Although networks and interorganizational relations have been mentioned by scholars in the strategy field over the past fifty years (Cyert & March, 1963; Granovetter, 1976; Pfeffer & Salancik, 1978), network theory did not receive much attention before the 1990s after Powell (1990). Today, network theory is recognized as a central part of explaining organizations' sustained competitive advantage. In the coming sections, we will review some of the most central work on network theory and its implications for organizations' innovation ability, more specifically by looking at Powell et al. (1996) and Uzzi (1997) and the path they made for subsequent research. Then, moving from the mid-nineties to recent studies in 2018, we cover different trends and the dynamic development of network theory research.

Powell et al. (1996) claimed that rapid technological development required more capabilities to succeed than a single firm could internally possess alone.

The locus of innovation arises in learning networks rather than in the individual firm (Powell et al., 1996). Firms without external ties are becoming rare. Often, they turn to their network for collaboration in joint ventures, alliances, or mergers. Powell et al. studied the pharmaceutical industry as they early understood the advantage of network effects. Further, firm age and historical growth no longer matter for engagement in external relationships. Growth is instead the outcome than a determinant of partnerships. Mainly resources, capabilities, and access are the criteria for inclusion. These findings were novel to prove why network theory is essential in explaining organizations' successful innovations.

While Powell et al. approach innovations as a result of knowledge exchange in networks, Uzzi (1997) argue for embedded relationships. Embeddedness and interfirm networks facilitate economies of time, integrative agreements, improvements in allocative efficiency, and complex adaptation while holding power to derail performance by making firms vulnerable to exogenous shocks or isolating them from information existing outside of their network (Uzzi, 1997). Vital characteristics were now social ties, the network structure, and the participants' positions within the network. Relational embeddedness was explored by later studies, such as Moran (2005) and Perry-Smith & Manucci (2017). The former tied it to innovations through the concept of social capital. That is, how the configuration and quality of the network ties explain performance. Structural embeddedness is suitable for routines, and relational embeddedness is better for innovation-oriented tasks. Managers that can cultivate intimate ties and draw on these relationships in uncertain times are a valuable asset. However, a combination of network closure and holes among the members is more optimal. An echo chamber creates no radical new ideas, nor do members standing a valley apart.

What we found to be interesting digressions from Leakey & Lewin and Tiger & Fox (1978; 2017) is how the unformalized rule of reciprocation creates a network of obligation among network participants. A web of indebtedness leads to clusters of interdependencies that bind individuals and organizations together into highly efficient units that transact with each other back and forth with more embeddedness as a result.

Gulati (2000) shows that the potential access to strategic information derived from embeddedness in quality networks leads to different strategic behavior of firms. It helps to have a network to scan to catalyst potential partner formation and alliances. Such strategic alliances are covered by Gulati et al. (2009, p. 20), Leiblein et al. (2002), and Rothaermel et al. (2004).

Dyer & Singh (1998) highlights network resources, such as interorganizational asset connectedness, partner scarcity, resource indivisibility (coevolution of capabilities), and the institutional environment. They push for cooperative advantage originating in the network's interorganizational ties, in contrast to firm-individual competitive advantage. While both Powell et al. (1996) and Uzzi (1997) directly approach network theory, Dyer & Singh (1998) indirectly recognize the value of interfirm networks through the relational view. Their focus is on the emergence of relational rents from knowledge transfer across dyads and networks.

Later, a stream of research aiming to address network ties, structural holes, and network closure emerged. Due to conflicting explanations in the literature, we have chosen to interpret the term «closure» as connected to dense networks or strong network ties and «structural holes» as sparse networks or weak network ties. Ahuja (2000) views the innovative functions of direct and indirect ties as facilitating resource transfer and information transfer. Structural holes lead to expanded information diversity, but malfeasance often arises as well, in sum leading to reduced innovation output. While structural holes may lead to the recombination of resources and novel ideas, unconnected people are tough to coordinate. In contrast, dense networks reduce barriers to coordination but lack the information variance to birth radical innovations, as Burt (2004) noticed. Obstfeld (2005) suggests third-party involvement to balance sparse and dense networks to optimize the network advantages. This balance relates to the strength of inter-firm ties, as described by Capaldo (2007) and Tiwana (2008). The former looks at a dual network architecture with the ability to integrate a large body of weak and strong ties to boost innovation. Capaldo (2007) mainly studied incremental innovative tasks. Tiwana (2008) focused more on radical innovation technology, describing developing alliance ambidexterity as a potential capability to overcome the issues. We find that the literature seems to agree that a combination of dense and sparse networks will yield the best innovation possibilities.

With Phelps (2010), we move to the alliance part of network theory, based on Dyer & Singh's (1998) relational view and cooperative advantage. It focuses on how alliances cultivate exploratory innovations. For example, in industries that are R&D heavy, the sharing of capabilities and know-how enables both partners to create something faster. Further, it highlights how both firms gain access to each other's networks through the alliance, creating a network of networks. Gnyawali & Park (2011) take it one step further by exploring the difficulties with coopetition and confirm Phelps's (2010) reasoning how alliances are still beneficial for the more radical innovations. When successful, cooperation often becomes a best practice in the industry, creating more alliances and leading to even more innovations. Bell (2005) finds the CEO's managerial network to enhance innovativeness. Furthermore, such informal friendships were found to be an important source of novel information, whereas institutional ties were not found to provide the same, suggesting the latter is mainly transmitting known information.

Lastly, Zhang & Li (2010) researched service intermediaries' (e.g., law firms, accounting firms, and so on) effect on the network. Given their overlapping roles for many of the network's members, they inhabit extensive knowledge of its inner workings while at the same time working in different networks, bringing in information variance, and serving as knowledge disseminators. The authors found that new ventures benefitted from the service intermediaries' extensive know-how of the industry's institutions and best practices, leading to faster growth. Eisenhardt et al. (2008) and Baum (2000) found similar results when small firms partner with established firms for innovation. As Carlile ((2004)) observed, it leads to knowledge transfer across firm boundaries.

The Ecological View

The ecological perspective, also known as the Ecological view or ecosystem theory, looks at the environment of organizations. It is an extension of Darwin's Evolutionary theory, viewing organizations as living organisms in an evolutionary cycle (Darwin, 2003). Evolution helps explain the rise and decline of firms and under which conditions some survive where others do not. The market is in a constant state of natural selection from competitive forces, giving firms harsh conditions they have to adapt to, leading the survivors to have a better fit. An industry's ecosystem is a good predictor of how the industry will develop over time. To use Malthus's theory of population density, we will have on the margin the number of organizations the ecosystem can support (Malthus, 2007). That is, access to resources predicts mortality and founding. One too many, and the whole population risks collapse. As Malthus argued in his book, two assumptions regulate his theory on human population size. First, Man needs food to exist. Secondly, the passion between the sexes is necessary and will remain nearly in its present state over time. We argue these assumptions can be transferred to organizations in the form of resources and capitalism's eternal pursuit of increasing wealth. As long as the environment offers opportunities, more organizations will be founded to take advantage. The opposite holds true for mortality.

Hannan & Freeman (1977) is one of the earliest advocates for ecological pressures on populations and try to explain organizational and industry structural variations by their environmental conditions. The environment will, over time, determine population composition in age, size, and organizational form, as well as dynamics around founding rates and mortality rates. This focus on mortality rates was followed by Carroll & Swaminathan (1992), who tied it to organizational form, leading them to discuss density dependence theory. The industry's density, legitimacy, and competition will strongly predict the direction and organizational mortality. In addition, the firm's environment will determine its configuration if it becomes a generalist or a specialist (niche) and its profits and longevity.

Brüderl & Schussler (1990), building upon Stinchcombe's (1965) liability of newness, discusses the liability of adolescence. They found that the risk of mortality is higher in the younger phase of the company life cycle. They found an inverted U-shape correlation and attributed this pattern to resource access and environmental forces, together with determining the probability for a firm's resources to endure past peak mortality into the mature phase.

The evolutionary processes were taken further by Aldrich & Ruef (2006), discussing organizations' natural selection cycle. The environment pushes forward the choice of organization form, degree of explorative or exploitative activities, the return on competitive or cooperative advantages, and if a winning formula could be sustained for more extended periods. The keywords are variation, selection, retention, and the struggle that together harbingers fit and survival.

In a more modern take, Hillman et al. (2009) try to explain how organizations wish to reduce environmental interdependence to combat uncertainty posed by their ecosystem. The degree of dependence on one's surroundings determines the power balance and, again, performance. To minimize uncertainties, firms either absorb dependencies through mergers, vertical integration, joint ventures, and other interorganizational relationships, with board appointments, political lobbying, and strategic executive successions, or firms innovate their way out of the dependency. Casciaro et al. (2005) find resource dependency as a powerful explanation of interorganizational action. Both Hillman et al. (2009) and Casciaro et al. (2005) build upon Pfeffer & Salancik's (2003) theories on external environmental constraints affecting organizations in different ways.

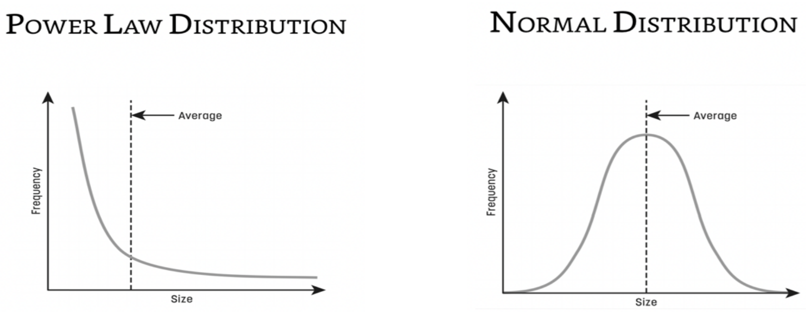
In a twist, Helfat & Raubitschek (2018) discuss how organizations can use environmental constraints to innovate value-capturing business models by creating and controlling their own platform-based ecosystems. This requires innovation of more dynamic capabilities, as Teece (2018; 1997) explains, which the firm may have developed from adapting to the existing forces in its environment.

Accelerator Programs

We will now cover what the literature says about accelerator programs. That extends to its participants, their functions, and how the network ties together. We begin with accelerator programs themselves, second by startups, followed by investors.

From Winston Smith & Hannigan (2015), we learn that early-stage entrepreneurial accelerators, with the most elite being *Y Combinator* and *TechStars*, take a small equity stake in the startups they accept into their programs in exchange for sponsor capital and the acceleration service. The accelerator programs are highly organized and have a strict application and selection process, including an on-boarding procedure (Cohen & Feld, 2011; B. Hallen et al., 2014). The most prominent accelerator programs are early-stage accelerators. They are often confused with *incubators*, programs that provide overlapping services to those of accelerators. Incubators are characterized by aiding individuals with ideas to incorporate their business and overcome some of the early hurdles of moving beyond the idea-generating phase and into more organized forms (Feld & Mendelson, 2019). Though we acknowledge the value incubators provide, not all startups use them, and they have less sway over the outcome compared to accelerator programs. Often, accelerators offer incubator services as well. Already incorporated startups seek out business accelerator programs with the intention of fast-tracking their growth. The accelerator provides a time-limited program consisting of mentorship, cohorts, and education, as well as access to its network of investors and sponsors (Feld & Mendelson, 2019). When the program ends, the survivability of the startups is better understood, and they participate in a public pitch, named demo day, to potential investors in order to finance their next phase of growth.

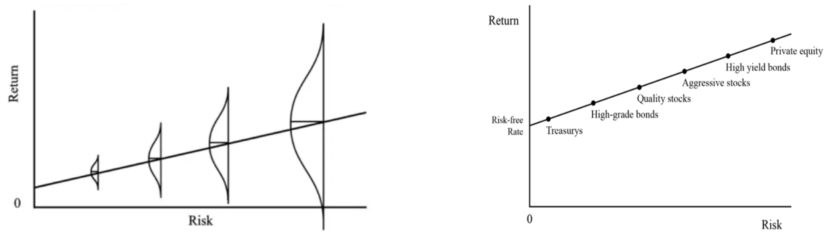
Combined, startups, investors, and accelerator programs try to innovate the future by allocating resources to the ideas they believe are worthy of survival. Their role in society is vital and unmatched. Their common denominator is their belief that change is the only constant in the world. Mallaby (2022) covers a concept named *the Power Law* which governs the entrepreneurial industry's outcome distributions and explains the participants' motivation to partake. The distribution is exponential and could be described as a more extreme version of the Pareto distribution.



Figures: Mallaby (2022) Power Law & Normal distribution

Unlike the traditional normal distribution of outcomes that govern the public stock market, the startup market is a market with more losers and a few super winners (Gompers et al., 2010). Only a small percentage of startups survive long enough to achieve extraordinary performance.

To place the startup outcome distribution on the risk-reward scale that governs modern financial theories and compare it to traditional investments, we learn that the prospect of higher returns typically involves added risk in the form of the variability of returns and increased potential for loss, placing such investments alone at the far end of the scale (Vishwanath, 2007). No other investment alternative offers the same intense risk-reward.



Figures: Outcome distribution along the risk-return line and the asset classes (Vishwanath, 2007).

Accelerator programs are in themselves network models and therefore fit well with our theoretical perspectives of network theory and the ecological view. We argue the industry's behavior as a whole is explained as well. Initially described by Von Neumann & Morgenstern (2007), further developed by Nash (1950), and later by Barnett & Hansen (1996), we have the aspect of games within the industry. Whenever one participant moves, the others adjust to that move to stay competitive. Accelerator programs enable faster innovation, creating a vacuum, which incentivizes founders to do more startup creation, and new investors to enter the industry, repeating with the result of increasing the amount of innovation going on in the ecosystem as a whole. We observe the Red Queen phenomenon and how it explains the industry's competitive forces. As explained by Barnett & Hansen (1996), The Red Queen hypothesis states that organisms must constantly adapt, evolve, and proliferate not merely to gain a reproductive advantage but also simply to survive while pitted against ever-evolving opposing organisms in an ever-changing environment. This point is further confirmed by Kim & Mauborgne (2005) that organizations find themselves running faster and faster just to stay in the same place. This is an ecology of organizational learning. In a sense, one can measure how good an idea is by the likely investment a competitor will make to undermine one's advantage. This thinking about our competitors' reactions is the cornerstone of game theory, scenario planning, and war game techniques, and how we assume the entrepreneurial industry will behave facing change as a group.

The advent of the entrepreneurial industry we know today was around the 1950s (Mallaby, 2022). Earlier business models were different and more focused on traditional manufacturing. A few large companies usually dominated each industry in a pattern explained by Coase (1937). Such models demanded different forms of organizing, and perhaps more importantly, a different form of financing. The regional banks were the typical lenders, and only in special cases larger national banks would be available. These conditions were not ideal for spurring entrepreneurial risk-taking. The advent of the semiconductor in the United States turned this trend around (Mallaby, 2022). A new type of founder and business model was conceived, and with it, traditional financing evolved. *Adventure capital*, as it was known, funded riskier bets with the potential to capture a large upside than traditional lending could. Arthur Rock is accredited to have pioneered venture capital (Mallaby, 2022). A new group of founders and investors believing in power-law distributions emerged. Today, we have a strong startup environment. This trend is, interestingly enough, still partly explained by Coase (1937) and Williamson (1991), where our form of organization has changed with a shift in sourcing behavior based on modern innovations, though we recognize that the venture capital network's role in combining the effects of the market and the corporation as quite unique. The accelerator program delivers to startups resources in the form of its network, contrasting with classical theories for organizing sourcing in-market and sourcing in-house. We now source it in the network.

The industry is in constant change. Over the years, unwilling regional bank lenders were replaced by willing adventure funds. With this, the power balance moved in favor of the capital owner, able to set unreasonable demands against the founders. Then the funds themselves started competing for startups, and the balance shifted to the idea owner. Out of this imbalance, the modern accelerator programs grew. Y Combinator was started by Paul Graham in 2005, fed up with traditional venture capital's treatment of founders. Today the gap is still there, constantly swinging, never staying for long in the equilibrium state. Negotiation power is highly dependent on the business cycle, creating booms and busts, the actions of other participants in the network, as well as the rules governing the ecosystem.

The investors participating in the accelerator program network share similar characteristics. They follow the power-law, which means they go after asymmetrical risk-reward bets that are probable to fail, where a handful of extreme winners would compensate them for a lifetime of losers (Mallaby, 2022). Their common goal is value creation, accomplished by sound investment analysis, strategic portfolio composition, and a timely investment entry and exit strategy (Feld & Mendelson, 2019). To better mitigate risk, different investor groups spread their exposure across several dimensions. Each type of investor specializes in a certain financing stage, follow-up investment strategy, number of deals they pursue, uses a typical financing type, has a preferred holding period, a specific industry they understand, a country they know, a governance style they apply, and so on. Importantly, they rarely invest alone and prefer to have investor syndications (B. L. Hallen, 2008; Hochberg et al., 2007, 2010; L. Zhang et al., 2017). Traditional financial quantitative models are not overly useful when a startup is too young to have numbers to modulate. The venture investor's investment analysis is therefore based on qualitative factors, emphasizing people and ideas, and inhabits greater perceived risk. Deviations from the past are these investors' hunting ground. We have several groups of startup investors: Family & friends, angel investors, seed investors, early-stage, mid-stage, late-stage, and enter at IPO. The groups correspond to the different financing stages a typical startup undergoes on its way to an initial public offering ('IPO'). Their money is unsurprisingly named the same way as startup money, seed money, and so on. The divisions are shown in the pyramid below.

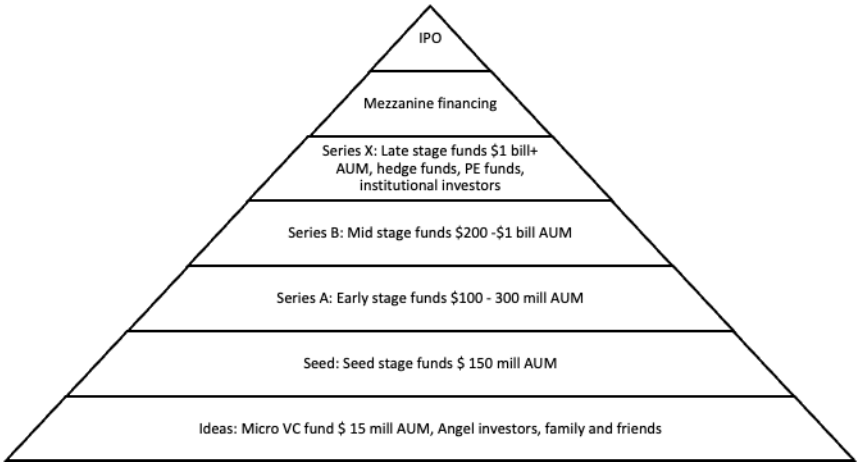


Figure: The Funding pyramid. Based on *Venture Deals Feld & Mendelson 2019*

Not yet mentioned are grant and award money. Not part of any particular financing stage, a grant comes from various business competitions set up with private, corporate, or public funds. The grant often represents a one-time opportunity for the startup, and the amount is relatively modest. In general, the grant does not involve equity in the startup. Rather, it is a prize won in startup competitions or granted as part of public support programs for entrepreneurs, with its real value being recognized in the industry.

The typical startup financing cycle is, as mentioned, broken down into multiple rounds (Cassar, 2004; N. Berger & F. Udell, 1998). The figure below shows a graphical representation.

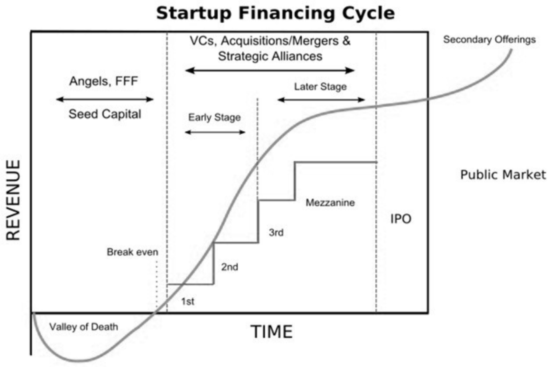


Figure: The startup financing cycle (https://www.theventure.city/blog/who-finances-the-startup-journey)

Before each stage, the startup produces the paperwork and makes it accessible for due diligence. Then the investors will send a contract, known as a *term sheet*. It includes a market value estimate, the number of shares they are willing to buy, what price they are willing to give, and a list of demands to protect them from excessive risk. In combination with the financing stages, we have different forms of demands. The investors take equity in the startups they back, but to mitigate the risk, they do rarely take normal shares in the company. Since the early stages translate to higher uncertainty, investors demand convertible preferred shares. Typically, they give the holder a liquidation preference, have an anti-dilution clause, redemption rights, board seats, veto rights, later financing round participation rights, and the right to convert to standard equity at will. This share class gives the investors better rights than the founders, which hold the standard equity.

Though the investors correspond to one of the stages, they carry all kinds of names. If we exclude the founders and friends & family, the traditional ones are angel investor, angel fund, venture capital fund, corporate venture capital fund, and private equity fund. At the top of the pyramid, larger institutional investors often enter when the probability of an IPO is higher. The role of investors is to finance startups in exchange for equity in their company. They adhere to the power rule of skewed risk-reward relationships, typically placing many small bets to create a portfolio of companies, then set out to cultivate them for success before they exit their investment.

The angel investor, or a syndicate of them, is typically defined as high-net individuals (Feld & Mendelson, 2019). Often, they made their fortune as entrepreneurs themselves. Their strategy is to seek out people they wish to back with their knowledge, network, experience, and money in a less formal way compared to other investor groups (DeGennaro & Dwyer, 2014; Kerr et al., 2014; Preston, 2004; Wiltbank & Boeker, 2007).

The venture capital fund is defined as a "fund committed to investing in early-stage companies" (Mallaby, 2022). A typical venture capital fund structure consists of three entities. The first is the management company owned by the senior partners, similar to how other partnerships work. Most funds have a limited time frame, and as they are started and retired, the management company lives on. The second entity is the general partnership. It is controlled by the senior partner that manages the funds. The third and last entity is the limited partnership or the actual fund that invests in startups. This vehicle is for the investors (the limited partners) and the portfolio companies. There is really no limit to how many such funds the senior partner can manage in parallel. The managers are raising money for their funds the same way that startup founders are: they initiate a new fund, set a specific mandate and timeframe, then venture out to potential investors for funding. As targets are discovered, they make investments. The next step is typically to assist the startup to enhance the fund's chances of a successful investment and exit down the line (Bernstein et al., 2016). This organizational structure is the current best practice in venture capital. To understand how VC firms invest, we first look at their timeframe. Notably, their commitment period (or investment period) and their investment term (their life span). The first governs the time a VC must find and invest in new companies. Since startups take years to mature and a fund has a defined life span, investing too late will impact the exit. Therefore, once the period is over, no new firms will enter the portfolio, they can only add more capital to existing holdings. To stay active, new funds must be started. In the investment term, the fund will participate in a new financing round for the startups they own. Typically, a fund has a time horizon of ten years and a commitment period of five years, with options to extend a few years should something happen to negatively impact the exit value.

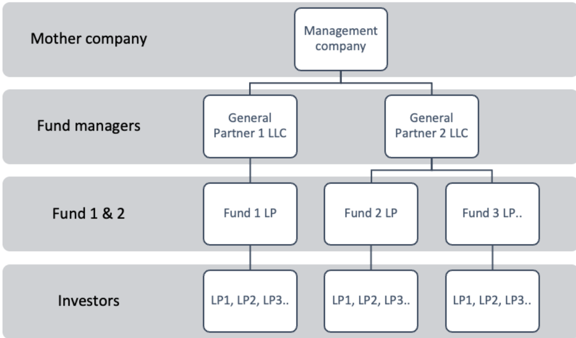


Figure: VC structure (Venture deals, Feld & Mendelson, 2019)

The corporate venture capital fund shares similarities with that of the venture capital fund, though they are a younger creation and have corporations behind them. Based on the success of venture capital funds, corporations created their own venture capital arms in the 1990s to access this part of the market (Feld & Mendelson, 2019). It is defined as “the investment of corporate funds directly in external start-up companies” (Chesbrough, 2002). This excludes corporate accelerators. In contrast, to venture capital funds, they hold innovation outcomes as well as financial measures as success metrics, leading to different behavior in holding periods and follow-on investments (Dushnitsky & Shapira, 2010). Their goals diverge, too, with the corporate venture fund having strategic objectives complementary to its parent, whereas the venture capital fund’s main goal is profits. This translates into performance incentives for partners in the latter, whereas the managers in the former are not allowed to take a performance fee or start multiple funds. Protecting the parent may cause conflicts with the startups over innovations, patents, and competition or translate into hostile acquisitions. This ruins the fiduciary duties the corporate venture fund managers have towards their startups, the founders, and their co-investors. Dushnitsjy & Shapira (2010) further found that incentives lead to the actions taken and again to the resulting performance or investment outcome. By not taking part in the upside, the results worsen. Chemmanur et al. (2014) find that corporate venture capital-backed firms are more innovative than venture capital-backed firms due to the effects of industry knowledge combined with a greater tolerance for failure, enhancing the startup's performance probability. Not to mistake them, corporate accelerator programs contrast with corporate venture capital (S. Winston Smith, 2021). The latter's mission is thought of as an external search for innovation materialized as a relationship between an established firm and a startup (Smith & Shah, 2013; Wadhwa & Kotha, 2006), where the former provides accelerator services. As Reuber et al. (2018) found, organizations looking for entrepreneurial ideas are part of the global ecosystem of opportunities and opportunity seekers, connecting in a network of relations. Dushnitsky et al. (2005), Benson et al. (2009)), Maula et al. (2012) and Winston Smith et al. (2013) talk about corporate venture capital as a strategic tool to access early-stage potentially disruptive ideas from the organization's ecosystem through structured equity relationships, pointing to its use for explorative innovation. Further, Basu et al. (2011) find that firms in industries with rapid technological change, high competition, and weak appropriability engage in greater corporate venture capital activity. Lee et al. (2018) find that CVC is used for exploitative innovations, too, though to a lesser degree than exploration. Limitations are connected to a lack of control and tougher governance, leading to geographical and industry constraints in such types of investing (S. Winston Smith, 2021). This further involves a lack of targeted innovation and shows characteristics of broader bets in an effort to quantify and hedge against a variety of uncertain outcomes. Corporations already represent a coalition of investors seeking to minimize individual risk (Coase, 1937; Williamson, 1991). Firms use several tools for risk management, and corporate venture capital is another similar instrument, placing multiple small bets on a set of uncertain future outcomes, assisting them in hedging against potentially disrupting radical innovations that will cause harm to their core business.

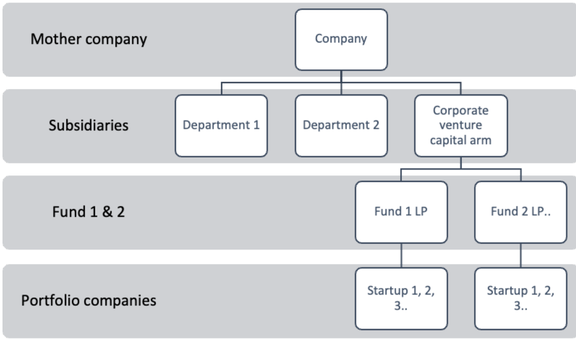


Figure: CVC structure (Venture deals, Feld & Mendelson, 2019)

Discussion of the literature

Having completed the literature review, we critically discuss what we think are the main findings in an effort to tie our theoretical perspectives and phenomena together. Then we will identify a knowledge gap to fill and show how our research question may fill the gap. Then we derive hypotheses.

In our review, we identified the most prominent literature from the knowledge stream on our theoretical perspectives and phenomenon. We learned how the machinery of the accelerator program network and all its components work. Based on what we have learned about the nature of accelerator programs, we view our choice of theoretical foundations as good lenses through which to interpret our phenomenon. From this exercise, it became apparent to us the fact that accelerator programs are still in their youth. Aware that academic research takes years to be produced, we realize we have a higher probability of discovering something novel writing on the subject of accelerator programs. Our Web of Science and Google Scholar searches uncovered far more articles on venture capital and corporate venture capital than on accelerators, as expected based on our Google Ngram search. Therefore, confident in our approach, we continue locating knowledge gaps adhering to our research question. In what follows, we extract what we perceive to be the biggest ideas from our literature search and compare them to our research question.

With Network Theory, Powell et al. (1996) big reasoning was the observation that networks increase the access to resources for their participants and subsequently their survivability and performance. As a trade-off, they had to expose themselves to dependencies. Uzzi (1997) observed that embeddedness in networks leads to larger knowledge exchange, meaning corporations were able to learn faster when they were close to others. The trade-off was the same, though to Uzzi (1997), it appeared as exposure to network shocks. Moran (2005) and Perry-Smith & Manucci (2017) found that the search for social capital leads companies to enter into relational embeddedness. Since the latter leads to more innovation, it also shapes how the network configuration looks, with closures and holes. The trade-off is the risk of echo chambers, where all but one are superfluous. Ahujja (2000) had confirmatory observations, arguing for a balance of structural holes to avoid group-think and distance-caused coordination issues. Gulati (2000) argues for the strategic advantage of network memberships, allowing participants to stay in "the know," spotting opportunities when they arise, which counter the trade-offs. Dyer & Singh (1998) makes a push for cooperative advantage derived from interfirm networks as a stronger force than a single company's ability to compete. The trade-off will be toll bridge rents from gatekeeper nodes' position in the network. Obstfeld (2005) adds the insight to use objective third parties to add balance, mitigating the disadvantages which were mixing itself into the network benefits. Tiwana (2008) backs it up. Phelps (2010) argues for the effect of a network of networks, where access to one opens access to another. It was observed to reduce time and increase innovation rates. Further, Phelps (2010) discovered the scaled economics of networks, where the big ones become bigger. Lastly, Zhang & Li (2010) observed how small companies gain advantages through interdependent networks with larger firms. Eisenhardt et al. (2008) and Baum (2000) found similar results. The exposure leads to growth.

With the Ecological View, we observe how the concept of evolution helps explain the rise and decline of organizations based on their fit with environmental conditions. Similarly, the total resource access in the environment explains the number of firms and the size of the networks through constraint conditions. Together, the environment will, over time, shape the networks. All participants have a strong dependency on their ecosystem. Brüderl & Schüssler (1990), together with Stinchcombe (1965), observed how younger firms struggle to survive compared with mature firms, contributing this to environmental constrained resource access. Aldrich & Ruef (2006) points to natural selection, where the ecosystem determines what organizational form survives. Hillman et al. (2009) make a case for environmental interdependencies and how firms strive to reduce them to minimize uncertainty. Lastly, Helfat & Raubitschek (2018) utilize the previous literature to point to the creation of stronger business models that are specially designed for a world where networks and ecosystems rule. They called them platform-based models, where a company builds and controls the ecosystem on which others operate.

We perceive the accelerator program to bridge most of the advantages and disadvantages brought up by the articles, overcome the concerns, and incorporate the strengths. An accelerator program offers a wide network where it connects small and large corporations around the globe. It acts as a gatekeeper node, central in the network, steering the resource and knowledge flows, creating dependencies, embeddedness, and connecting structural holes.

Based on Powell et al. (1996), we view accelerator programs as access to resources that increase survivability for young firms. That low barrier access allows them to form embedded relationships with others in order to get much-needed know-how. Uzzi (1997) points to exogenous shocks and echo chamber thinking as the price to pay for admission, though we argue startups have nothing to lose, all to gain by staying in the know. Given their lack of bargaining power, young firms should try to create a cooperative advantage together as a larger entity. That requires interorganizational ties, which an accelerator program offers them. As Dyer & Singh (1998) points out, it may not be free from relational rents. Usually, the accelerator program demands a small portion of equity. We take Ahuja (2000) position and argue that unconnected people are tough to coordinate and paying a membership fee for access to a network seems to be worth it. Given the power-law, a majority of the equity paid becomes worthless anyways and for the winners it gets dwarfed compared to the growth in their wealth. Obstfeld (2005) and Tiwana (2008) suggest third-party involvement to balance any network. We argue the accelerator program could be the objective part, adding variance and pushing best practices at the same time to optimize innovation, balancing the startup network and the investor network, and working with aligned incentives to create win-win situations for all participants. In short, creating a network of networks, the way Phelps (2010) describes. We remember that scaled economics shared is a powerful model and we argue accelerator programs are ideal for utilizing it. By becoming larger, their ability to connect to new networks grow, leading to increased attractiveness, attracting new networks to gravitate towards them, and so on in an increasing expansion. The bigger they get, the more startups and investors they help, creating a growing number of win-win situations. Lastly, the venture capital industry history points to a need for a balance of power. If accelerator programs can educate founders on how to swim with sharks without getting eaten, startups might grow fast without taking on excessive risk.

Based on Helfat & Raubitschek (2018), we consider the similarities between accelerator programs and platform models. Creating and controlling a platform allows them to set the rules for the participants. Appreciating Hannan &

Freeman (1977), simply by setting the constraints, ecological pressures will shape participants. In the case of startups, accelerator programs recreate evolutionary pressures where they learn how to fail faster and adapt to become better. As young firms often fail because of the lack of resources (Brüderl & Schussler, 1990; Stinchcombe, 1965), the accelerator program keeps them afloat at a critical stage and gives them access to growth capital from investors, raising the probability of survival until a more mature stage. It creates an arena where founders can struggle well, giving them a fair shot at success. Further, the Malthusian principles can be transferred to govern the harsh environment of startups. Any population is constrained to grow no faster than the pace of the resource growth in its environment. In their small beginnings the number of societal resources funneled towards startups and founders determine how small or large the startup population will become. It cannot break through the resource barrier by itself. We argue that an accelerator program model is a form of organization of resources from society with the sole aim of increasing the Malthusian population limits. Lastly, we call on Darwin. Constraints lead to the survival of the fittest startups or those with the most productive use of resources. We see no reason that accelerator programs teaching away best practices to all with the result that more startups survive through the needle eye of evolution should not be in the best interest of society.

As we have discussed, we think our choice of theoretical perspectives and phenomena fits well together to explain firm survivability and performance.

Knowledge gap

Some factors surrounding the accelerator program seem to us to not be covered in the existing literature we have combed over. Namely, how the participants in the accelerator program network affect the network itself. Most literature covers what the network can do for the participants' performance and not what the participants can do for the network's performance. It typically covers other types of industries and less within the energy industry. This is the knowledge gap we have identified. We initially asked, *Which network factors yield increased performance for an energy industry accelerator program?* With that as our question, we aim to cover the literature gap with further research and analysis.

Hypotheses

Hypotheses are tentative answers to research questions and are expected but unconfirmed relationships between the variables (Straits & Singleton, 2018). Based on our research question, we derive the following hypotheses to uncover network performance-related factors. They are built on the format of predictors and predicted effects. Since the network consequences are tough to measure directly, we use appropriate proxy variables to help us measure them indirectly. We intend to test them to get answers in order to plug the knowledge gap. We draw our hypothesis from a variety of sources, including most prominent theory and literature, data observation, and intuition and guesses.

Network access reveals how the networks connect to vital input factors. Such as ideas and capital, which are critical for accelerator program performance. This will illuminate the absorptive ability of the network to take up founders' ideas and investors' capital, to create innovative startups.

A: If there are many unique investors and/or a large investor distance, then we expect to observe more startups in the network and an increased total sum raised compared to networks with fewer investors and/or smaller investor distance.

When the resource access increases, we expect it to increase performance. Therefore we expect large investors to attract and deploy more capital.

B: If higher than the mean investor AUM, then we expect to find investors deploying more capital and/or more financing rounds, measured by the increased sum raised.

Second, we look at *network attractiveness* and the willingness to participate in the network. We contemplate if certain characteristics are present the network will attract more participants.

C: If there are many startups and/or a large startup distance, then we expect to observe more investors in the network and an increased total sum raised compared to networks with fewer startups and/or smaller startup distance.

Third, is *network scalability* and network growth over time. The accelerator program should allow for scaled network advantages. We think we will observe that annual growth numbers do not return to the mean, rather the larger accelerator networks only get bigger with time.

D: If the accelerator network has a comparable large participant cumulative aggregate growth rate, we expect to observe that scaled network advantages apply, and that the fastest grower becomes the largest total network.

Fourth is *network embeddedness,* with network clustering and spread. We contemplate clustering may increase homogeneity and suppress the goal of innovation. Heterogeneity and variance are often associated with new impulses and a flow of ideas.

E: If we observe many unique countries present in the network, then we expect to observe more innovation and experience, measured by patents, startup age, and investor age.

F: If the number of unique investor countries present in an accelerator portfolio is high, we expect a larger number of investors to be attracted.

Fifth, we wish to observe the combination of *network closure and holes*. A common measure of knowledge transfer in the literature is innovation, where patents are used as the proxy (S. W. Winston Smith, 2019). For innovation, a balance between closure and holes is preferred to avoid echo chambers and valleys. We expect that the best startups have an increasing amount of patents to protect their market share and reduce competition. Further, investors seek predictability and are willing to give startups with patents more capital. We think these measures show the relationship between closures and holes.

G: If we observe a large number of patents, we expect to find a larger amount of total sum raised.

H: If we observe a large number of patents, we expect to find a larger number of investors in the accelerator program.

Sixth is *network centrality*. Since betweenness centrality may describe nodes that are tollbooth bridge nodes between other nodes, we expect central nodes to exhort a gravitational pull that increases their centrality. It may have a network scaling effect.

I: If the accelerator network has many participants relative to the time it spent acquiring them, then we expect to observe a better betweenness centrality, measured by total network participants growth.

Seventh is *network on network access* where new connections lead to an increased number of participants and improved value of the network. Seeking out a better investor group, with long experience and more capital, could add to performance. Such investors are more likely to have increased syndication across multiple investor networks. The same goes for attracting investors from a large number of countries relative to only having domestic investors.

J: If a greater multi-network interaction by investors, then we expect to observe higher than mean investor AUM and/or higher than mean investor age.

Eight is the *network construct of small worlds*. We think we will find the same importance in attracting the right startups as with investors. They become tight networks, or small worlds, when they are placed in cohorts in the accelerator program. Their inherent qualities will therefore affect the cohort through spill-over effects. We think their age and previously undergone accelerator programs affect their quality.

K: If a young and un-accelerated startup joins the accelerator program, then we expect to observe better performance than with an older, previously accelerator startup, measured by total raised capital.

3. Methodology

Research design

Our thesis is set up as descriptive quantitative research, where we seek to collect and analyze quantifiable data to say something about our phenomenon. Our chosen research method is called *network development research* on the interorganizational level (Carpenter et al., 2012). That implies that we study causes that affect a network, in which the focal actors are organizations that combine to form a network. We try to recognize patterns and determinants of network formation and change, from the network and non-network constructs working as *predictors* (causes). Therefore, particular network features become the *predicted effects* (consequences). Our study is non-experimental and observational, producing descriptive results and leading us to draw conclusions from the data we have. To analyze the causes and consequences, and the patterns, we perform an *association study* (Altman & Krzywinski, 2015; Borgatti et al., 2009; Breiger, 2004; Carrington et al., 2005; Crossley, 2010). We utilize research techniques obtained from *Social research: Approaches and Fundamentals* by Straits, B. C., & Singleton, R. A. (2018). Our research is designed in such a way that other researchers should be able to repeat it.

We wish to briefly explain our choice of method. We initially planned and performed multivariate linear regressions as the main analysis, only to discover it would not fit our data in any meaningful way. We have added our regression outputs in the appendix. We recognize the importance of choosing a type of analysis that fits with the collected data. From econometrics, we know that basing network analysis on observational data to examine causal relationships potentially leads us to biased estimations when using ordinary least squares ('OLS') (Stock & Watson, 2020). Several articles cover similar issues on how network sampling frequently suffers from inaccuracies and biases (Erickson & Nosanchuk, 1983; Granovetter, 1976). As found by Wooldridge (2002) and Krackhardt (1988), autocorrelation might bias OLS regressions and generate unreliable significance test results. Given our sample size, how our network data was structured, and the use of a non-random sampling method, problems occurred in analyzing the data, and we changed our method of analysis.

Our new method became *association*, where we discover similar patterns across networks and utilize triangulation to interpret our findings. Though we cannot deduct statistically significant causality without performing regressions, our findings are definite insights about our units of analysis based on discovered patterns. We argue that our method derives useful knowledge about causes that affect our accelerator program networks. Our sample size covers a niche population and generalizability across populations would not be statistically defended by the multivariate regression method either, given the small sample size, so we feel confident in our new method. Making our findings generalizable may be a fruitful area for future research.

Measurement and sampling

The units of analysis are the entities under study (Straits & Singleton, 2018). Our unit of analysis will be three energy accelerator networks, comprising the Industry partner accelerator, the Dedicated company accelerator, and the Dedicated industry accelerator. To represent them, in order, we have found Equinor & TechStars Energy Accelerator ('Techstars Equinor'), Shell Gamechanger, and TechStars Alabama Energy Tech Accelerator ('Techstars Alabama'). We will analyze the accelerators, the startups passing through them, and the investors and sponsors surrounding them. Each accelerator represents a slightly different style. TechStars Equinor is a *partner-with*, where a corporation works in alliance with an industry-specific accelerator to develop startups the company could have the advantage of. TechStars Alabama is more of a *stand-alone*, where they are independent in their selection of startups. Lastly, Shell Gamechanger is a *corporate* one, where they select and accelerate startups in-house they believe could add value to the company.

Our operational definition for the target population is all the entities that fulfill the membership defining rule (Straits & Singleton, 2018). In this study it is all participants in the aforementioned accelerator program networks. That means all the startups, the investors, and the sponsors attached to the programs from their founding till today. Our target population and sampling frame are therefore the same. We are not handpicking our sample, we are sampling the whole population. Given our small population and use of software programs for analysis this was manageable and we thought it would produce more reliable and valid results, adding less bias. Namely, sampling error, omitted sample bias, and investigator bias in the selection and data analysis. The need for a normally distributed sample is, therefore, less important and dealt with during data inspection and cleaning. Given our operation definition, we will not be able to say anything conclusive about similar populations or generalize our findings across populations. To confirm, more comparable research would have to be performed, opening up an area where future research may test our associations with other energy accelerator programs. In the population, we observed 3 accelerator programs with a total of 70 startups and 362 unique investors. This gave us an N = 432.

Reliability and validity

Reliability is concerned with questions of stability and consistency, whereas validity refers to the congruence or «goodness of fit» between an operational definition and the concept it is purported to measure (Straits & Singleton, 2018). Based on our quantitative approach, drawing the whole population in, we expect to find the same results in repeated trials, avoiding the ecological fallacy. When it comes to validity, our method fits well to describe parts of the reality of the population we are looking at, though the method and small sample hinder our findings from being transferable to other populations without the same conditions. Naturally, the problem with omitted variable bias, that some extraneous variables hold explanatory power, is tough to entirely remove.

Data collection

Our study is limited to the time interval from the founding of the programs until the latest available data. The observations were discovered in a sequential manner. We began with our three accelerator programs, then mapped all their accelerated startups, then mapped all the investors connected to each of the startups. This allowed us to create networks of startups and their investors, using the accelerator program as the center node. The population and sample included N = 432 observations. Further, we used the literature and general knowledge to create preliminary variables that could explain network factors and performance. All our data was collected during the timeframe March 1st and March 30th in 2022, representing the latest publicly available Q4 data from FY 2021, but not Q1 FY 2022 or later. It will be a cross-sequential study since we are looking at multiple points in time. Our sample consists of private and public companies operating in a regulated industry, requiring them to send data to the government, where third parties like us can access it as reports and financial statements. The data is retrievable, reliable, delivered in a standardized and comparable format, and checked by auditors. We have utilized a database named Pitchbook for this information. Pitchbook gathers these financial reports and makes them accessible through search and standard formatting. We have done cross-checking of random data samples, using CB Insights, CrunchBase, and annual reports from company websites to confirm the accuracy of the Pitchbook data. Our data is mainly secondary data collected, which is often the case with quantitative analysis. Secondary data has multiple strengths over primary data, and we do not see the distinction between primary and secondary data as impacting our findings.

Ethical considerations

As researchers we have responsibilities. General Data Protection Regulation ('GDPR') regulates the ownership of the information we use, limitations to its use, and requires consent. Our chosen data is publicly available data, meaning we will not infringe any privacy or break anonymity. Whenever public information about individuals was accessible, we strived to preserve their anonymity or chose not to pursue it. Further, we believe it is our ethical obligation to ensure the validity of our findings and true objectivity. With our research design and collection process, we have strived to behave ethically and morally correct. If others should deem our work worthy of inclusion in the knowledge stream, we must ensure the highest quality. When in doubt, we looked to guidelines for research ethics (Straits & Singleton, 2018), the Navigation Wheel (Kvalnes, 2019), and The Norwegian Center for Research Data.

Data processing

Based on the industry data, we manually built a novel numerical dataset in a computer-readable form (Appendix). The data set is accompanied by a descriptive table with more comprehensive information about the data points and variables, such as their construction and potential deviations from the main data collection (Appendix). We had to codify several variables before we entered them, translating the text into numbers. This process helped us control the data quality as we went along. We have dedicated a lot of time to data entry to avoid errors and other fat-finger problems.

We ended up with 50 variables that we thought held explanatory power. The variables are categorical variables and continuous variables, covering geographical data, time data, financial data, and more. The finished data matrix had the shape of 432 x 50, giving us 21.600 data points. Limited by data availability our data matrix has a larger focus on the startup observations, and we lack information on some investor observations, leading to missing values and zero values. This reduced the total number of quality data points, making statistical regression tough, and eventually led us to pursue association testing.

With data exploration, the goal is to get a clear picture of the data to determine appropriate statistical analysis and data modifications (Straits & Singleton, 2018). We used a set of basic data science and analytics libraries from Python to perform data inspection, cleaning, and exploration (Package description, Appendix). Those were Numpy (Oliphant, 2015), Pandas (McKinney, 2010; Reback et al., 2020), Matplotlib (Hunter, 2007), Scikit-learn (Pedregosa et al., 2018), and Statsmodels (Seabold & Perktold, 2010). Together they mimic the functions of the statistical software R, another popular research tool. We remember that extraneous variables may pose rival explanations to the relationship between our variables, decreasing the accuracy of our conclusions. Therefore, it was vital to work out if our data with a high degree of certainty could help us test our hypotheses. We did a univariate analysis where we examined each variable with tables, graphs, plots, and standard measures. Lesser variables were fixed, replaced, or removed altogether. It helped us reduce zero inputs and handle missing values. The latter was handled on a variable basis without auto-filling. These cells are flagged in the data set. Wild-code checking and consistency checking was performed. Further, we plotted the whole variable set out and frequency distributions were used. Our work helped us curb potential multicollinearity in later analysis. Though we had some, we did not remove outliers. Our literature search told us we would find extreme outliers due to the power-law. Our interpretation leads us to believe removal might be a larger source of error than inclusion. We became attentive to 7 observations with extreme age that we contemplated removing later on should an analysis warrant it to avoid skewing the data. Those analyses would then be based on 63 observations. Removing based on age instead of other variables allowed us to keep important power-law distributions in the models.

Instrument reliability is the way of ensuring that any instrument used for measuring experimental variables gives the same results every time (Straits & Singleton, 2018). Given that we use Tableau and Python to conduct a quantitative study on secondary data based on our whole population, we will not experience any reliability or validity complications originating from our instrument. To be thorough, we performed Intercode Reliability testing, where both of us performed the calculations and matched our answers.

Analysis

We do network development research (Carpenter & Li, 2012). That implies we look for causes that affect the network. Calculating bivariate associations is not sufficient to test causal hypotheses and we have to use a multivariate association analysis where we include several predictors. To model the relationships we seek out which independent variables affect the dependent variable to uncover patterns in the data and build models. Since network consequences are tough to observe directly, we utilize appropriate proxy variables to measure effects indirectly. Avoiding reciprocal causation, noncausal association, indirect effects, and omitted variable influence is a priority. We used a data visualization program, Tableau, to generate network maps to visually map network structures from our data set. We use geolocational data to map the accelerator actors, then apply other variables to filter them out to model our expected patterns. Further, based on our data matrix we performed calculations to retrieve associative patterns.

Dependent variables

We have chosen strategic dependent variables, or predicted effects, based on our expected patterns. Our variables are continuous. We have selected the variables we think will be the best proxy variables to pick up the network effects from the predictors. A variable may serve as a predictor in one model and a predicted effect in another. They will not appear in the same model, to avoid causing confusion in the results.

First, *Unique_startups,* a variable containing the number of startups. Second, *Total_raised,* a variable holding how much funding each startup has been able to raise through its life cycle. Third, *Number_investors,* shows the number of investors per startup. Fourth, *Network_size,* contains the accumulated number of participants in the network. Fifth, *Patents,* describe the connection between startups and their patents. Sixth, *Startup_age,* portrays how old the startups are. Seventh, *Investor_age,* shows how old the investors are. Eight, *Scaling_effect,* a variable possessing the relationship between the growth of accelerator network participants and time. Ninth, *MeanAUM_investor,* carrying the assets under management for each investor.

Independent variables

We have chosen strategic independent variables, or predictors, we think will influence our expected patterns. To assist us in this process, we utilized our Python data exploration and did several rounds of experimentation (Code script, Appendix). We used a Pearson's R correlation matrix (Appendix) to identify better predictors. We have mainly used continuous variables as independent variables, though we have categorical ones, so-called 'dummies', to filter out the needed observations in each model. This is practical when we analyze individual networks.

First, *Number_investors,* shows the number of investors per startup. Second, *Investor_startup_distance,* represents the kilometers between investors and their startups. Third, *Investor_accelerator_distance,* represents the kilometers between investors and their accelerator. Fourth, *Startup_accelerator_distance,* represents the kilometers between startups and their accelerator. Fifth, *Unique_startups,* a variable containing the number of startups. Sixth, *Accelerator_CAGR,* a number presenting growth rates of participants over time. Seventh, *Countries,* a variable holding all the unique countries connected to our individual observations. Eight, *Patents,* describing the connection between startups and their patents. Ninth, *Multinetwork_investors,* a variable showing which investors are present in more than one accelerator network. Tenth, *Startup_age,* portrays how old the startups are. Eleventh, *Times_in_accelerator,* holding the number of times each startup has partaken in an accelerator program. Twelfth, *MeanAUM_investor,* carrying the assets under management for each investor. Thirteenth, *Investor_area,* describes the total geographical area the investor networks covers based on its perimeter nodes. Fourteenth, *Startup_area,* describes the total geographical area the startup networks covers based on its perimeter nodes. Fifteenth, *Investor_age,* describes how old the investors are.

Models

To analyze the relationships we expect to find in the data, we have designed the following models. Our research and hands-on data experience lead us to create them this way. The need for control variables was replaced by our use of a triangulation approach. The models are presented in table form to assist the reader in the linkage between predictors and predicted effects on network factors through proxies. Our analyses findings will be presented in the next part.

Model	Network performance consequence	Proxy consequence	Predictor(s)
A1	Network access	*Unique_startups*	*Unique_investors + Investor_distance*
A2	Network access	*Total_raised*	*Unique_investors + Investor_distance*
B1	Network access	*Total_raised*	*MeanAUM_Investor*
C1	Network attractiveness	*Unique_investors*	*Unique_startups + Startup_distance*
C2	Network attractiveness	*Total_raised*	*Unique_startups + Startup_distance*
D1	Network scalability	*Network_size*	*Accelerator_CAGR*
E1	Embeddedness	*Patents*	*Countries*
E2	Embeddedness	*Startup_age*	*Countries*
E3	Embeddedness	*Investor_age*	*Countries*
F1	Embeddedness	*Unique_investors*	*Countries*
G1	Network closure and holes	*Total_raised*	*Patents*
H1	Network closure and holes	*Unique_investors*	*Patents*
I1	Network centrality	*Scaling_effect*	*Accelerator_CAGR*
J1	Network on network	*MeanAUM_Investor*	*Multinetwork_investors*
J2	Network on network	*Multinetwork_investors*	*Investor_age*
K1	Network construct small worlds	*Total_raised*	*Startup_age + Times_in_accelerator*

4. Analysis and findings

We will now discuss our analyses and extract findings. We have performed multivariate analysis through calculations and network mapping to observe whether a number of independent variables and a dependent variable have an association that confirms our theoretical expectations.

Data matrix

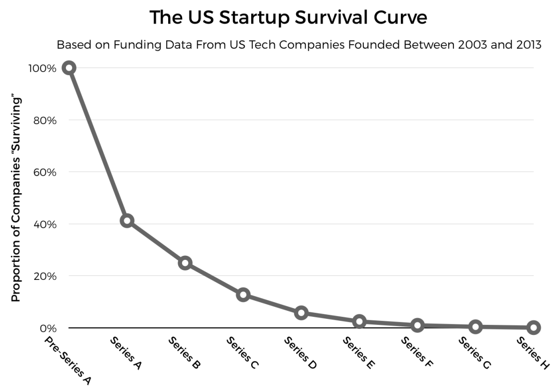


Figure: Average survival rates US startups (Techcrunch: https://techcrunch.com/2017/05/17/heres-how-likely-your-startup-is-to-get-acquired-at-any-stage/?guccounter=1)

From the academic literature we found the power-law distribution dominates industry data. From the figure above we observe it in action for the general startup survival rate across the US market in the period 2003 to 2013. Observe that a mere 0,01% survived until their last financing round before exit. We think it speaks to the validity of our data that we found similar distribution patterns all over our sample data. From the age of participants (average: 7, median: 6 max: 70), capital raised ($mill: average: 13, median: 1 max: 233), financing rounds completed (average: 4 median: 3 max: 15) investors attracted per startup (average: 7,5 median: 4 max: 46), investor asset size ($mill: average: 16594 median: 180 max: 725000), patents (average: 1,6 median: 0 max: 21), and even in the country distribution which is heavily centered in the USA (total: 58%, startups: 54%, investors: 60%). The Gaussian bell-curve is nowhere to be found in these key metrics. Our chosen variable distributions are shown as histograms in the Appendix.

Equity financing was almost ubiquitously the only form of financing conducted within our dataset (98,5%). It comes in the form of seed, angel, or early-stage financing rounds. Some utilize the newer phenomenon of equity crowdfunding, but that was a minor observation (5,7%). We saw few reaching as far as the last financing rounds. We had few exits in our data set and then mainly by bankruptcy (4,29%) or acquisition (7,14%). A majority had never been accelerated before, though the extreme value was 8 times, which increased the average to 2. The average reached 4 financing rounds, with the extreme being 15 rounds. The average value of capital raised was 13,2, with the median being 0.85, and the max 233. We found the investor base to consist of individuals, angels, venture capital funds, corporate venture capital funds, private equity funds, corporations, and more. The industry distribution was naturally heavily skewed to the energy industries for startups and investors.

Hypothesis testing

We did association analysis and triangulation to uncover relationships in our data. We define the criteria for keeping or rejecting our hypothesis as the presence of a directional pattern of correlation that repeats in the observations of all subsamples. We are able to determine if a plausible relationship among variables exists in our sample, though we cannot rule out whether it is due to chance factors or extraneous variables, hindering the extrapolation of our findings across populations and their predictive powers. Whereas only a few figures are strategically included in the text to add value to our explanations, all our maps, figures, and calculations are available in the appendix. All analyses are derived from data found in our data set.

Network access

Network access reveals how the networks connect to vital input factors. The knowledge can be used to optimize resource dependencies in the network. To uncover such factors' impact on accelerator performance, we will triangulate our analyses from the three models A1, A2, and B1, to answer our connected hypotheses.

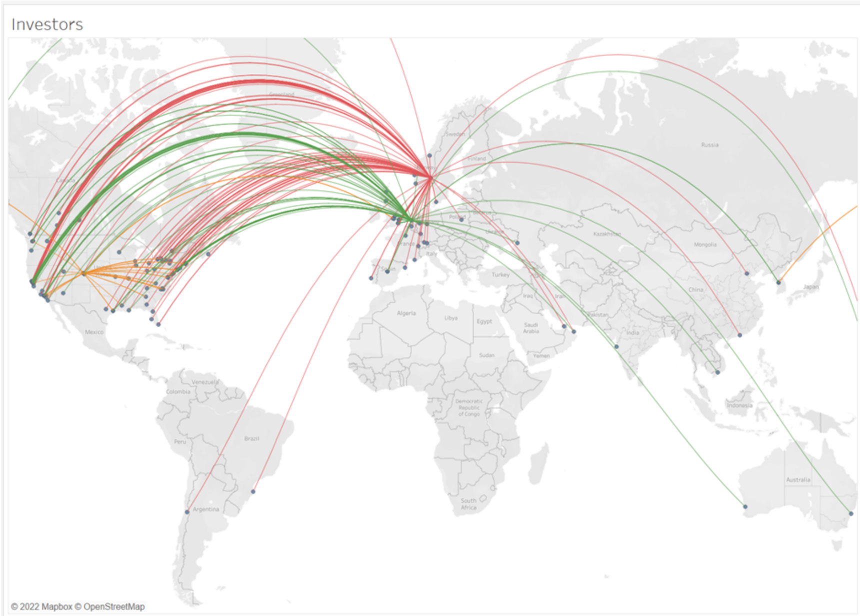


Figure: Investor networks (Green: Shell Gamechanger. Orange: Techstars Alabama. Red: Techstars Equinor.).

The network map above and the subsequent analyses for A1, A2, and B1 are based on 258 of 362 investors. We omitted values due to unavailable location data for angel investors and undisclosed investors.

The findings from Model A1 describe how the investor network and its geographical distance from the accelerator relate to the startup network. In Figure 1 (Appendix), we learn that the Techstars Alabama's 57 investors have an average distance of 2658 kilometers with a total of 103.642 kilometers, covering 19 startups. The Shell Gamechanger has 160 investors, with 6259 kilometers on average, and 718.748 kilometers in total, for 22 startups. While Techstars Equinor holds 187 investors, averages 5834 kilometers, and totals 845.863 kilometers, to control 29 startups. Techstars Alabama's network is easily the most geographically condensed one with the lowest number of participants. Further, we observe that the geographically larger network access has more investors and startups.

Model A2's findings show how the investor network and its geographical distance from the accelerator relate to capital raised. Figure 2 (Appendix) displays that the Shell Gamechanger network was able to raise $336 million with an average distance of 6,250 km, Techstars Equinor $285 million with an average distance of 5,834 km, and Techstars Alabama $143 million with an average distance of 2,658 km. Again, the larger geographical investor networks stick out. However, when looking at the distance in absolute terms it does not seem to be a clear tendency. What this can allude to is that accelerator proximity and closeness to investors do not seem to be a limiting factor when it comes to raising capital.

We remember our hypothesis A: *If there are many unique investors and/or a large investor distance, then we expect to observe more startups in the network and an increased total sum raised compared to networks with fewer investors and/or smaller investor distance.* Based on our findings, we conclude that this statement holds.

The findings from model B1 portray how another feature of the investor network, their assets under management, appears to affect the total capital raised. As figure 3 (Appendix) shows, the networks with more potential firepower do not show any clear tendency to raise more capital. Techstars Alabama while having the highest mean AUM of $35,307 million has the lowest total capital raised of $143 million, Shell Gamechanger with a mean AUM of $574 million between that of Techstars Alabama and Techstars Equinor of $141 million has the highest total capital raised of $336 million. Though it does seem to be a slight indication that the higher mean AUM of the investors leads to more money being raised per startup per unique investor, with Techstars Alabama raising $2,5 million per additional average investor, Shell Gamechanger 2,1 million, and Techstars $1,5 million respectively.

For hypothesis *B: If higher than the mean investor AUM, then we expect to find investors deploying more capital and/or more financing rounds, measured by the increased sum raised.* We found this to partially hold. While we did not see this relationship in our data on an aggregate level there seems to be a relation between the amount invested and capital raised on an investor average level. Brought together we see a relationship between investor network coverage in terms of attracting investors and raising capital. However we do not find a clear indication that investor assets seem to be related to more funds raised for the accelerators. We concluded that it does not hold.

Network attractiveness

Network attractiveness talks about the willingness to participate. The insights can be employed to become attractive for target groups to boost expansion. We will triangulate the analysis from models C1 and C2.

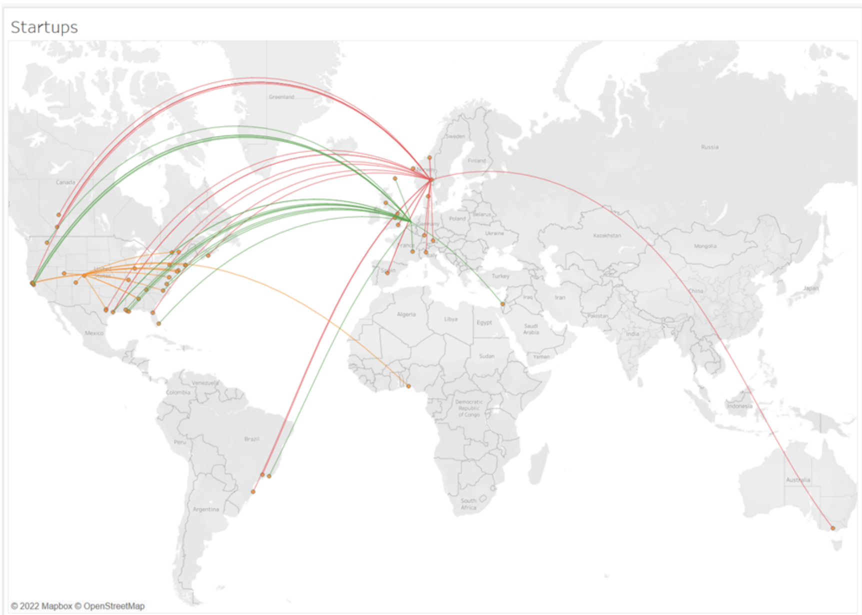


Figure: Startup networks (Green: Shell Gamechanger. Orange: Techstars Alabama. Red: Techstars Equinor.).

The network map above and the subsequent analyses in C1 and C2 are based on 70 startups and their locational data. The reason the total number of investors per accelerator below exceeds the number of individual investors is due to the double-counting of multi-network investors.

The findings from Model C1 talk about to what degree startups' geographical distance from the accelerator attracts more investors. From figure 4 (Appendix) we learn that Techstars Alabama has 19 startups with an average of 2230 kilometers, a total of 42.366 kilometers, attracting 57 investors. Shell Gamechanger has 22 startups, an average of 5931 kilometers, a total of 130.488 kilometers, for their 160 investors. Lastly, Techstars Equinor has 29 startups with 4824 kilometers as the average, and 139.907 kilometers as the total, drawing 187 investors. We discover that the larger startup networks have attracted more investors.

From model C2 we learn to what degree startups' geographical distance from the accelerator attracts more capital. From figure 5 (Appendix), we find that Shell Gamechanger was able to raise $336 million with an average distance of 5,931 km, Techstars Equinor $285 million with an average distance of 4,824 km, and Techstars Alabama $143 million with an average distance of 2,230 km. This translates into a pattern where the larger startup networks raise the most capital. It is noteworthy that Techstars Equinor with its accelerator model has managed to attract more startups and investors both in absolute numbers but also in terms of the number of countries.

Our hypothesis *C* was*: If there are many startups and/or a large startup distance, then we expect to observe more investors in the network and an increased total sum raised compared to networks with fewer startups and/or smaller startup distance.* Based on our findings we can confirm that our hypothesized relationship is found in the data.

Network scalability

Network scale relates to network growth behavior over time. Knowledge about growth rates might incentivize the accelerator programs to aggressively pursue expansion to get big first and dominate the industry. To uncover such performance factors, we will triangulate the analysis from model D1.

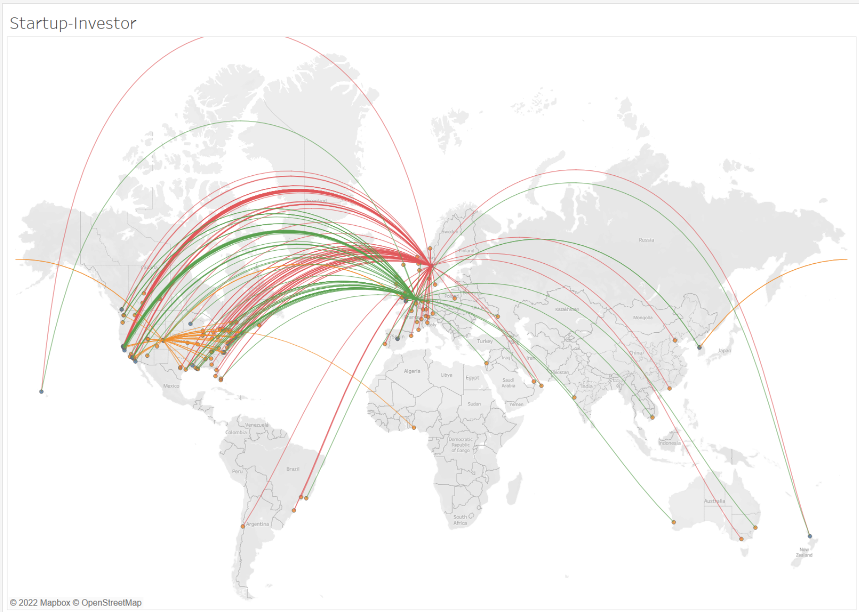


Figure: Total network participants (Green: Shell Gamechanger. Orange: Techstars Alabama. Red: Techstars Equinor.).

The network map above and the subsequent analysis in D1 is based on 70 startups and 258 of 362 investors. We omitted values due to unavailable location data.

Our findings from Model D1 seek to explain if networks with high participant growth rates are likely to become the largest network. Figure 6 (Appendix) tells us that Shell Gamechanger on average adds 1 startup and 7 investors annually, Techstars Alabama adds 6 startups and 25 investors, while Techstars Equinor adds 6 startups and 37 investors. Compared to their years in operation, Shell Gamechanger's total growth over 26 years has been 22 startups and 160 investors, a record it took Techstars Equinor 5 years to beat, totaling 29 startups and 187 investors. Techstars Alabama growth over 3 years resulted in

19 startups and 57 investors. With their current growth, they will probably pass Shell Gamechanger in a few years. We find that accelerator program age has less to say for network size compared with participant growth, which points to scalability. Given that both Techstars Accelerators (Alabama and Equinor) are able to outpace Shell Gamechanger in such a short time further showcases that accelerator programs are able to effectively leverage established networks, and achieve network on network effect. As mentioned in C, Techstars Equinor in particular outperformed Shell in network attractiveness. When adding the time dimension from D we observe they outperform them on pace both in relative and absolute terms.

Our hypothesis *D* states*: If the accelerator network has a comparable large participant cumulative aggregate growth rate, we expect to observe that scaled network advantages apply, and that the fastest grower becomes the largest total network.* Our findings do support this relationship.

Network embeddedness

Network embeddedness speaks about node clustering and spread. Insights into the relationship between embeddedness and performance may lead accelerator programs to seek out clusters or a country expansion strategy to add variance. We will triangulate the analysis from the models E1, E2, E3, and F1.

Figure: Geospread patents (Green: Shell Gamechanger. Orange: Techstars Alabama. Red: Techstars Equinor.).

The network map above and the subsequent analyses E1, E2, E3, and F1, are based on all 70 startups.

From Model E1 we seek out the relationship between countries and patents, and how increasing the number of countries and the additional variance new countries represent may lead to more innovation measured by patents. From figure 7 (Appendix) we observe that Shell Gamechanger's startups hold 32 patents and their network includes a total of 17 unique countries, giving us a ratio of 1.88 . Techstars Equinor's network holds 44 patents and spans 22 unique countries, resulting in a ratio of 2. Lastly, Techstars Alabama's numbers are 5 patents and 6 countries, giving us the ratio 0.83. We observe that adding new countries is associated with additional patents for the accelerator program and implies increased innovation.

Model E2 illustrates the link between startup countries and the age of the startup network. We ask if geographical spread adds experience and knowledge to the network, measured through startup age. From figure 8 (Appendix) we find with Shell Gamechanger that it spans 3 unique countries and averages 7 years of age. Techstars Equinor covers 8 countries and has 6 as the average startup age. Lastly, Techstars Alabama has 3 countries and an average age of 3.5 years. We do not find that the average age seems to grow with the number of countries for startups.

With Model E3 we analyze if the number of investor countries relates to the investor network age. Again we ask if geographical spread adds to the network or not. Figure 9 (Appendix) tells us that Shell Gamechanger's investors are on average 22 years old, while covering 15 countries. Techstars Equinor's investors are 13 years on average, spanning 16 countries. Lastly, Techstars Alabama investors are 24 years old and present in 4 countries. We do not find that the average age seems to grow with the number of countries. The total age however does seem to be indicative of the total number of countries and the total distance.

Our hypothesis *E* goes as follows*: If we observe many unique countries present in the network, then we expect to observe more innovation and experience, measured by patents, startup age, and investor age.* Our findings point to increases in patents with countries, though not in startup age or investor age. Therefore, we conclude this hypothesis is inaccurate.

Lastly, we analyzed Model F1 to find whether the increasing number of investor countries leads us to observe more investors. The reasoning goes that geographical clustering depresses the number of investors while variance inflates the number. Figure 10 (Appendix) shows that Techstars Equinor has 187 investors in 21 countries, while Shell Gamechanger has 160 investors in 15 countries, and Techstars Alabama in contrast has 57 investors in 4 countries. In ratio terms, the first has 8.9, the second 10.8, and the third 14.25 times as many unique investors as countries. In other words, additional countries are associated with additional countries.

For hypothesis *F* we had*: If the number of unique investor countries present in an accelerator portfolio is high, we expect a larger number of investors to be attracted.* We find this statement to hold.

The network embeddedness hypothesis test for our three accelerators speaks about node clustering and spread and seemed overall to give validity to it. We find that geographical spread is an overall good predictor for our data. Though we have to point out that both average distance and total distance proved to be better universal predictors than the number of unique countries within each accelerator network.

Network closure and holes

Network closer and holes point to echo chambers and valleys. Knowledge about how the network is shaped in that respect aids in efforts to increase knowledge flows and subsequently innovation. To uncover network closure and hole factors that affect the accelerator program performance, we will triangulate the analysis from the models G1 and H1.

Findings from our Model G1 intend to uncover the link between patents and capital raised. The reasoning is that patents proxy knowledge transfer and network innovation. From figure 11 (Appendix) we see that Shell Gamechanger's network holds 32 patents and has raised $336 million, Techstars Equinor's network has 44 patents and $285 million, and lastly, Techstars Alabama's network holds 5 patents and raised $143 million. For each extra patent, the first raised $11 million, the second $5.65 million, and the third $28 million. We find that the number of patents do not necessarily point to the amount of capital raised. We do find however on the aggregate level the 20 startups with patents in our sample captured approximately 69% of the total capital raised, or $530 million in total, and an average of $26.5 million, against those 43 startups without patents which captured approximately 31%, or $234 million, and an average of $5.4 million.

We remember our hypothesis *G: If we observe a large number of patents, we expect to find a larger amount of total sum raised.* Overall our data does support his statement.

In our Model H1 we describe the relationship between patents and investors. Financial theory states that patents reduce investment risk by protecting the product and therefore the revenue stream, leading us to think the investor network is closer around startups with patents. From figure 12 (Appendix) we observe that Techstars Equinor has 187 investors and 44 patents, while Shell Gamechanger has 159 investors and 32 patents, and Techstars Alabama in contrast has 57 investors and 5 patents. In ratio terms, the first has 0.23, the second 0.20, and the third 0.087 patents per investor. We observe a relationship between increasing patents and investors. This is supported on the aggregated level with the average number of investors per startup observed to be 2,5 times higher for those with patents compared to those without.

For our hypothesis *H* we have*: If we observe a large number of patents, we expect to find a larger number of investors in the accelerator program.* We found this statement to hold for our accelerator programs.

Network centrality

Network centrality tells us about the existence of nodes that bridge other nodes together. In doing so, they control interdependencies. Knowledge about such network nodes or the ability to become one holds value for accelerator programs. We intend to illuminate centrality with the triangulation of model I1 results.

Findings from Model I1 show us the relation between rapid network participants' growth and time. A node that increases through traffic exponentially rather than linearly might be a gatekeeper or a bridge. From figure 13 (Appendix) we observe that Techstars Equinor's network has an average annual network growth of 43 participants and a compound annual growth rate of 23%. Techstars Alabama's network has an average annual growth of 25 participants and a compound annual growth rate of 51%. Lastly, the Shell Gamechanger network increases participants with 7 each year and has a compound annual growth rate of 22%. We find that the networks are scaling exponentially and might become, or already be, central nodes in their network.

We had that *I: If the accelerator network has many participants relative to the time it spent acquiring them, then we expect to observe a better betweenness centrality, measured by total network participants growth.* We conclude that this statement holds.

Network on network

The network-on-network effect is a phenomenon where new connections lead to an increased number of participants and improves the value of the network. Knowledge of how to increase connections may enhance performance for the accelerator programs. To uncover such factors we will triangulate the analysis from the models J1 and J2.

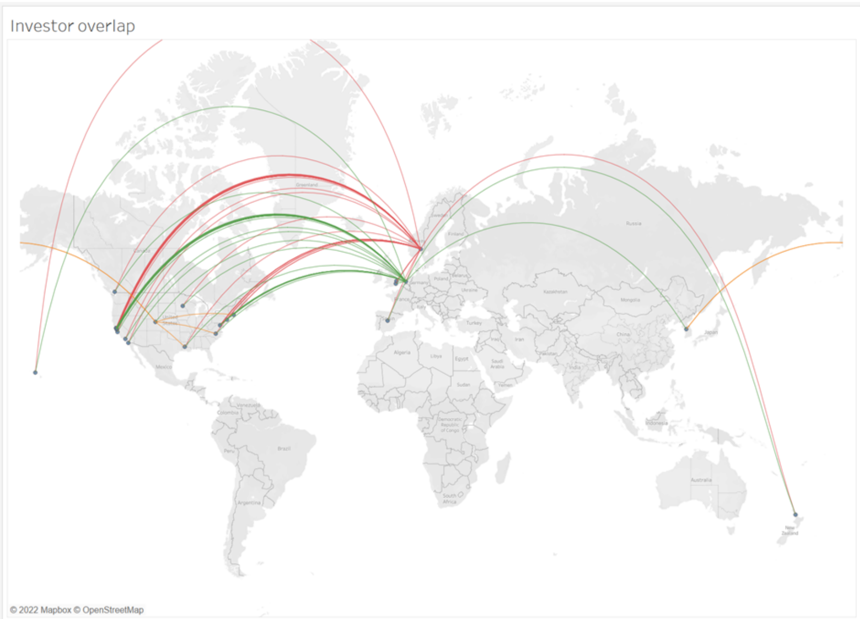


Figure: Investors that connect to multiple accelerators (Green: Shell Gamechanger. Orange: Techstars Alabama. Red: Techstars Equinor.).

The network map above and subsequent analyses in J1 and J2 are based on 258 of 362 investors. We omitted values due to unavailable location data.

From our Model J1 we observe if the multi-networked investors have more assets under management. The reasoning is that experienced investors will be welcomed into several new syndications given their previous network connections, capital funds, and status, in hopes of spill-over effects on younger investors. From figure 14 (Appendix) we observe that a mere 40 investors out of the 362 total span across more than one accelerator network. Of the 40, only 15 of them are investors with available data. Only 4 of them are in the third and fourth quartile of all 258 investors' assets under management. The other 11 are below, with insufficient or lesser funds. We do not see a systematic pattern where the investors with large funds are consistently invested across multiple networks. We speculate these networks to be competitor networks not willing to grant access to each other or are otherwise in competition over resources not suitable for sharing. Both capital from their investors and quality startups to invest in are scarce.

In our Model J2 we study the link between investor age and the status of being part of more than one investor network. The thought is to examine if experience from age leads investors to be welcomed into new syndications. In figure 15 (Appendix) we learn that only 5 multi-network investors are above mean investor age, whereas 35 of them are below mean age. The pattern shows us no clear linkage between high age and multi-network connections. Instead, it reveals that a majority of them are young. The investors share commonalities with the startups. They are often young and lack sufficient funding.

Our hypothesis *J* was that: *If a greater multi-network interaction by investors, then we expect to observe higher than mean investor AUM and/or higher than* *mean investor age.* Based on our analysis we conclude that this hypothesis does not hold.

Network construct of small worlds

A small-world network is one where the shortest path average distance is low. Knowledge about small network constructs helps accelerator programs build higher-quality startup cohorts. We will triangulate the analysis from model K1.

Findings in Model K1 reveal the relationship between a startup's experience, measured by age and the times it has been through an acceleration program, with capital raised. The intent is to observe if pre-accelerator program experience makes them better or worse. From figure 16 (Appendix) we observe that Shell Gamechanger's startup network has an average age of 7, average times in accelerator 1.9, and average raised capital is $21 million. Techstars Alabama's startup network has an average age of 3.5, average times accelerated at 1.5, and average raised at $7,5 million. Lastly, for Techstars Equinor the numbers are 6 years, 2.5 rounds, and $10 million. Based on our findings, we argue that an increased startup experience systematically leads the startups to raise more capital. Naturally, the survivors raise more in subsequent rounds as other startups fail and are eliminated from further financing rounds.

Our hypothesis *K* is the following*: If a young and un-accelerated startup joins the accelerator program, then we expect to observe better performance than with an older, previously accelerator startup, measured by total raised capital.* Based on our analysis, this statement does not hold. We observed the opposite. This finding tells us that while age is not the penultimate deciding factor we see that Shell Gamechanger has a higher average raise than the rest of the "above average age" category. There is a clear overall preference with startups above the average age both in terms of amount and number of investments made.

Ranking

Having concluded our analyses and answered the hypotheses, we will rank the three accelerator program networks. Our aim is to reveal which best practices other energy accelerator networks should consider assimilating to increase their probabilities for superior performance. The below table can be viewed as a scoreboard of our findings from above.

Network factor	Hypotheses	Effect
Network access	True	Strong
Network attractiveness	True	Strong
Network scalability	True	Strong
Network embeddedness	True/False	Weak
Network closure and holes	True	Mediocre
Netowork centrality	True	Strong
Network on network	False	Mediocre
Network constructs	False	Strong

Figure: Hypotheses condensed result overview

Network access proved to have an overall effect on the performance of an accelerator. It had an apparent effect on the number of participants in the network, though its effect on the total raised capital is somewhat more ambiguous (which is also reflected in the predictors used). Network attractiveness corroborates our findings from network access. The measure of distance (a proxy for network reach) seems to be a clear indicator of an accelerator's ability and reputation to attract new members. The measured effect on total raised capital however proves that it is a less clear indicator when it comes to performance. From our analysis we find strong support for network scalability. Techstar has a large preexisting network and presence, and both of its accelerator programs outperform the one from Shell regarding growth rates and overall dynamism. The effects of network embeddedness while present are not unidirectional. We observed clustering in certain countries, though we faced uncertainty when we tried to extract relationships from it. There seems to be a distinct effect from network closure and holes, even though model G1 was only proved on an aggregate level. Patents have been proven to be a great predictor of many performance metrics. We measure strong betweenness centrality in our network centrality analysis. This highlights yet again the strong expansion performance of the two accelerator types having a dedicated accelerator, such as Techstars, over that of one serving as a supporting function to a main company, such as the Shell one. When exploring for stratification of investors we did not observe any clear network on network effect, quite the opposite in fact. This could interestingly be due to intense rivalry or because younger investor models are more adaptable. While we found that our hypothesis for network construct of small worlds was disproven, we found conversely that its predictor for startup experience, age, to be a strong predictor of performance.

Globally we find that most of the predictors have a conforming effect on either the aggregate level or the accelerator level. We interestingly notice that the ones that are less clear are usually related to the total sum raised. Given that age seems to be a related factor, it is not surprising that there is a certain amount of equivocacy surrounding tests related to total age given we have accelerator programs of three different ages. Among our observed accelerator programs we do not find one obvious winner. There are many network factors that combine to create superior network performance. In our view, performance should be judged based on several factors. The first is enveloped by total members and their geographical spread combined. The second includes an accelerator's ability to generate value through its program. Third and last, is the rate at which it achieves the former two, its growth rate. In terms of the first performance factor mentioned above, Techstars Equinor, our IPA, was resoundingly the winner of this category, attracting numerous startups and investors from a large geographic area. In regard to the second factor, for its ability to generate value Shell Gamechanger, our DCA, had the strongest overall performance, where it raised considerably more capital than other accelerators, and on average managed to generate the most patents per startup. Techstars Alabama, our DIA, while ranking lower on the above metrics had a phenomenal expansion rate clearly outpacing the others in its achieved growth. Given its young age it will be interesting to see if it manages to keep up this dynamism. After deliberation we veered towards picking Techstars Equinor, with the industry partner accelerator model, as the overall best performer of the three. Given that we find the time factor to be positively leading to most other performance metrics, the fact that it has managed to outclass Shell which is substantially older in our overall performance metric is quite astounding. This is further elevated by the fact that Techstars Equinor had strong results in terms of generating value and managed to generate the most patents. It continues to show strong growth despite having the largest network, pointing to scaled benefits.

5. Discussion, critique, implications

We intend to compare our findings to the existing literature, explain their significance, before we tie it all back to our research question and answer it. Further, we will point out shortcomings in our research, inconsistencies and anomalies, before we suggest improvements to the research design. Next, we will place the study into a broader perspective by mentioning theoretical and practical implications. Lastly, we discuss areas for future research.

Interpretation of findings & Literature comparison

In our first hypothesis A we were concerned about the implications for an accelerator network should its investors be many and far geographically distant from its location. We contemplated the consequences to be mainly positive based on Powell (1990) and the resource access a larger network typically represents. To an accelerator program, ideas and capital are input factors. From our study we found both that an increased number of investors and an increased investor distance lead the networks to observe more startup participation (i.e., ideas) and capital raised. Related literature findings, such as Powell et al. (1996), describe that added connections result in access to extra capabilities and resources, resulting in a learning network, or an innovative network. We find no gaps or surprises compared to existing literature.

With our hypothesis B we considered that greater investor assets might lead them to deploy more of it with the consequence that the accelerator network's startups observe more capital raised in addition to extra financing rounds. We thought the consequence for the network to be positive based on Powell et al. (1996) and Eisenhardt et al. (2008) which showed us opportunities presented by networks would be seized by the larger firms. We found that not to be true in our data. This represents a gap with the existing literature.

For hypothesis C where we pondered the impact on an accelerator network should its startups be many and far from its location. We mused the consequences to materialize as increased investor participation and more capital raised, based on Powell (1990) and that any larger network often brings resource access. Our findings indicate that relationship to be true. Found by Eisenhardt et al. (2008) and Baum (2000), as long as the network offers opportunities, firms will be found to take advantage. We find no gaps or surprises compared to existing literature.

In our hypothesis D we wanted to study if accelerator networks scale as other network models are frequently found to do. We got the idea from Phelps (2010) and Dyer & Singh (1998) that found alliance networks cultivate faster innovations and consequently that successful alliances lead to more cooperation, creating a larger network. We found this pattern to hold in our data. The big networks tend to become bigger in a non-linear fashion. We find no gaps or surprises compared to existing literature.

With our hypothesis E we wished to uncover the relationship between clustering and innovation. Based on Uzzi (1997) we thought to find that geographical embeddedness isolated startups from new information resulting in a less innovative accelerator network. We found that more countries in the network increased the number of patents, a proxy for innovation, but was not however linked to increased investor and startup age. These findings are partly consistent with Uzzi (1997) and Moran (2005) that found network embeddedness and which networks one is a part of to help explain innovation performance. Compared to existing literature we find that our proxy for experience, age, has little association with country spread and clustering. This is therefore a gap compared to the literature.

In our hypothesis F we wanted to analyze if large investor networks or syndications have a gravitational pull on outside investors. We suspected this association to hold based on Gulati (2000) who stated embeddedness in quality networks is sought after for the access it provides to strategic information. We found this to be accurate. More investor countries and networks pull new investors in. Similar was found by Gulati (2009), Leiblein et al. (2002), and Rothaermel et al. (2004). We find no gaps or surprises compared to existing literature.

For hypotheses G and H we sought to uncover the relationship between accelerator network innovation, structural holes, and network closure. Based on Ahuja (2000), we suspected that innovation, proxied by patents, was found to increase when closures and holes were balanced, resulting in more investors and capital. Our thinking was that networks with close ties experience echo chamber tendencies while those with loose ties get coordination issues and that the balanced ones produce more innovation (patents) which again attract risk-averse investor capital. We found no systematic pattern in the data for each accelerator program leading us to think patents increase investor presence or capital raised. However, on an aggregated level we found this to hold. We now see our findings in line with what has been mentioned in the existing literature we have combed over and have no gap.

With hypothesis I, we look for tollbooth nodes and try to identify how an accelerator network can increase its network centrality to become one. We got the idea from the relational rents Dyer & Singh (1998) found can be derived from interfirm knowledge transfer. We found that historical rapid participant growth tends to continue for accelerator networks rather than decrease over time, leading us to view them as becoming more central in their network as they age. We find no gaps or surprises compared to what we would expect based on existing literature.

In our hypothesis J we talked about network-on-network performance, anticipating a connection where large and experienced investors pursued a multinetwork strategy. We based it on Zhang & Li (2010) who discussed that network participants with roles in overlapping networks bring with them variance and function as information propagators. This led us to think experienced investors would be welcomed into new syndications where both parties would benefit from the connection. We found that investors with substantial assets under management and considerable age were not in multiple networks. Though we found the opposite, that younger less capitalized investors sought out multi-networks. This gap surprised us, and we do not believe it is stated in the literature we have examined.

Lastly, with hypothesis K we reasoned startup experience, proxied through age and times underwent accelerator programs, would be a subtractor to accelerator network performance. Previous experience may signal inertia, inhibiting the assimilation of best practices, and disturbing the rest of the startup cohort. Our view comes from Hannan & Freeman (1977) who discussed how experience shapes firms and adapts them to a special environment, reducing their ability to change when faced with a new one. We found the opposite in our data, that startups with increased experience prior to acceleration systematically raised more capital. This gap surprised us, and we do not believe it is stated directly the same way in the literature we have inspected.

Implications

Having finished our research we have found some discrepancies with both theoretical and practical implications that we wish to elaborate on.

When considering our findings' impact on *existing theory and research*, our first reasoning is the enhanced explanatory power we found when marrying our two theoretical perspectives to understand the patterns in our data. Ecology theory added another dimension to Network theory and helped explain how the networks came to be, why we have competing networks in the first place and what they compete about, and network path dependencies. Merely applying Network theory makes us expect geographical clustering of investors and startups. Though we actually observed that the presence of accelerator programs increased the distance between them. The participants are still clustered, rather in new forms, aided by the ease of monitoring brought by the digital age. A shock to the system that to us is explainable by their collected ecosystem. Our second reasoning is similar to the first. Heterogeneity was thought to be geographically dependent as well, where greater distances introduce variance. We found that networks connecting to networks across the globe may add heterogeneity independent of geographical distance or closeness. Further, explaining it through network node distance is in our view not good enough. We observed an accelerator program with great kilometer reach that only operated in one country, which means theoretically the network was perceived as distant, but practically it was closer to homogeneous. Our third reasoning comes from the observation that the power-law distribution affects investors as well. Young investors inhabit many of the same characteristics as startups, searching for their own investors' capital to represent and again deliver performance by investing it properly. Their outcomes are non-linear and we observe extreme winners. In our fourth reasoning, we argue that research should incorporate new ways of measuring innovation and value creation besides patents. We observed companies with no patents, though much capital was raised. Fifth and lastly, we found that investors with substantial assets under management and considerable age were not in multiple networks. We reason this to be explained by ecology and how the networks compete against each other.

Reflecting on our findings' impact on *practice and policy* for accelerator programs, energy accelerator programs, and public policymakers, we have the following takes. First, what an appropriate accelerator target is should be reconsidered and not limited to startups alone. We observed established firms entering accelerator programs as part of turnarounds and directional strategy changes. We argue mature as well as young firms may benefit from acceleration in the form of value creation and innovation. Further, it may present mature firms with an option to private equity firm buyout when a need to redirect their focus and resources surfaces. Lastly, their success rate from acceleration may be comparably higher than startups considering their inherent experience and existing resources. We found that firms with increased experience prior to acceleration systematically raised more capital. Second, we reason industry cluster formation should be organized for. We found in our data that geographical clusters are popular, and their companies thrive compared to isolated companies. Third, following our observation on patents, we urge policymakers to simplify patenting processes to secure innovations and correct value appropriation for founders and investors. Though we did not find that patents are systematically connected to capital raised, we worry the lack of them may lead to misappropriations.

Bridging Network Theory and Ecological View

Given our findings, we propose for researchers to union or partially bridge Network Theory and the Ecological View more often when studying the phenomenon of accelerator programs. In our view, ecosystem understanding is supportive of networks. Typically, when researching local and global associations our proposed two-fold approach may yield greater insights into the mechanisms of networks by considering their influence from surrounding environments delivering rules and shocks. Said simpler, the network explains the behavior of participants and the ecosystem sets the rules that form the networks' interactions. We view networks as living networks, similar to living organisms, which interact in a way determined by their collective environment. From our study we found differences when comparing regional industry networks around the globe. We expected them to be similar and contribute their differences to their regional environments.

Limitations

We recognize that our study has weaknesses that influence the strength of our findings. These are something we tried to change but in the end, could not.

Since we do not intend for our readers to read more into our research than is warranted, we wish to discuss our data limitations and analytical shortcomings.

Generalizability across populations, such as other industry accelerators and other similar networks and ecosystems, is our main one. Though we have selected the leading accelerators in the energy industry globally and have all reasons to assume our findings could be used to say something about other similar accelerators, we cannot say it for sure. Our findings are representative of the sampled group only and cannot be assumed generalizable across populations. The correlation-causation troubles in analyzing our findings will stand out as a big limitation. We found a lot of correlations among our variables but we cannot draw inferences about observed behavior in the data. The real reason behind the observed behavior may be something else not covered in our data or the patterns are formed due to randomness.

Our analytical approach consisted of drawing associations among variables to form patterns in the data. Weaknesses with this approach and how it fits with our data and tools used could create weaknesses in our findings. Perhaps our phenomenon is better researched through a qualitative approach when applying association.

For our target population, we chose the leading accelerators in the energy industry in the Western hemisphere, introducing a Eurocentric view. Such a view limits the applicability of our findings to the Eastern hemisphere. Our study being culturally less diverse also implies narrower findings than if the opposite was the case.

Survivorship bias is inevitable when looking at past data of startups. Only the best survive long enough to be accelerated and surface in our data, while the power-law dictates that only the greatest survive until an exit, shaping how we measure performance. We cannot observe or say anything useful about the startups that never reach the accelerator stage other than the not useful advice that they should really try to survive and enter an accelerator. We do not suffer from sample bias, though we do have a sort of selection bias in the fact that we chose a small and specific population from the energy industry. If our study is to be replicated, we recommend choosing accelerators with larger populations or simply including a larger number of accelerators in the target population. Preferable across industries.

We repeat that we found all our data in a single database. Though Pitchbook is regarded as an industry leader and widely acknowledged, we cannot help but think that inclusion of multiple databases might have aided our initial research on the topic, pivoting our investigation to presently unknown areas of the accelerator industry, proving to be a value-add. Further, we suggest it will help when gathering the data and creating variables. Taken together, we think these steps will open for statistical regression analysis and improve the generalizability across populations, making the findings even more useful.

Our dependent variables were chosen based on theory to reflect the known influential network factors. We think they are well chosen. The same goes for our choice of independent variables. Though we are open to the possibility other unknown to us variables could hold significant explanatory power. Furthermore, we use independent variables as proxies for effects we are not able to observe directly. Within each proxy lays the fallacy that we are measuring other effects or the wrong one altogether. Should we change our predictors we might end up with different findings. Omitted variable bias is further a trap hard to avoid completely.

We derive our insights from a simplified model of reality, whereas reality is far more complex. From this follows another common source of limitations to all studies which lie in the assumptions used by researchers and the assumptions in articles they base their study on. Findings are only as good as the assumptions holding them together. Researcher bias is another certainty. It sneaks in and shapes our study. Other researchers will make their mark if they choose to redo our research. Another consequence is confirmation bias. As researchers, we always run the possibility we subconsciously have sought out data that confirms our initial suspicions and constructed data sets such that the analysis picks up our existing beliefs, confirming our hypotheses, rather than challenging them to a proper degree with true relationships. Naturally, such a bias is both limiting and tough to eliminate. Here the Man-with-a-hammer syndrome comes to play. We are limited by the tools we know how to use, which may hinder us in observing reality how it actually is and consequently choosing another route. The tendency to twist the problem to fit with our models will inevitably affect the solution.

Improvements to research design

From our post-mortem exercise, we have uncovered improvements to the research design that might advance the significance of our study and findings.
First, a larger sample size to unlock the possibility of choosing statistical regression rather than association. We still stick by the triangulation method regardless of regressions. Second, more data points on investors and startups. Third, considering the benefits of a qualitative dimension with industry interviews to the quantitative one. It may help better understand motivations unreadable from numbers and time data, such as the how and why network participants come together. Fourth, additional data sources to reduce zero inputs and missing data.

Future research

When conducting our study we came across surprises and uncovered new insights that we think are interesting and worthy of sharing. First, should our study be performed again, we suggest updating the research design to include our improvements. Second, hypotheses should be formed to answer the gaps we uncovered between our findings and the existing literature. Third, a longer time series study to observe the accelerator programs while they age would be appropriate. It would allow the study of weights of importance from knowledge, capital, and ideas, to long term network success. Fourth, performing our study with different predictors to see if the results are the same, as well as using our predictors across populations to measure their transferability.

6. Conclusion

We will present a summarizing conclusion of our thesis. Afterwards we will give recommendations to energy accelerator programs to which network factors the accelerators should prioritize to enhance their performance, based on rankings and findings.

Our phenomenon was *network effects on accelerator programs* and how it influences the mechanics of modern innovation practices. From that we asked *which network factors yield increased performance for an energy industry accelerator program*? We confirmed our question's relevance through literature gap spotting. From a decade of growing popularity, accelerator programs are now in high demand. We formulated several hypotheses rooted in literature to answer our main question. They suggested logical relationships that would affect network factors and lead to changed performance for accelerator programs. We found most of the expected associations to be in our data, leading us to conclude that network factors impact accelerator program performance. An accelerator network's access to ideas and capital, namely startups and investors, affects its performance. The network attractiveness draws in startups and investors and makes them willing to participate. In an increasing manner too, as network size creates gravitational pull. As the network grows, it reinforces embeddedness among its participants. Network closures and holes balance to determine the innovativeness and value creation of the network. As a third party, accelerators can serve as a balancing node to bridge holes, resulting in its increased centrality and importance in the network. Its position opens for new connections made through other networks, extending its reach further. With relational power follows the ability to construct the network to facilitate for optimal growth. We find that the network business model is strong. Initially, we uncovered a knowledge gap in the literature and designed a study to discover new knowledge to fill it. With our findings we argue we have answered our research question and filled the gap.

However, answers contain within them seeds of new questions and we have therefore proposed areas for future research endeavors.

Recommendations

Based on our findings, we have ranked the three programs against each other to unveil best practices that we argue should be assimilated by other energy accelerator programs to increase the probability for superior performance. We list them as clear and actionable recommendations.

- Redefine startup and include more experienced firms in the program
- Attract more startups in general and create larger cohorts
- Attract more investors in general, not just experienced ones with considerable assets
- Be open for startups and investors from all around the world to increase the probability of innovation
- Actively encourage patent protection
- Choose a dedicated model and not a general model when establishing an accelerator program
- Actively research and assimilate best practices in the accelerator industry

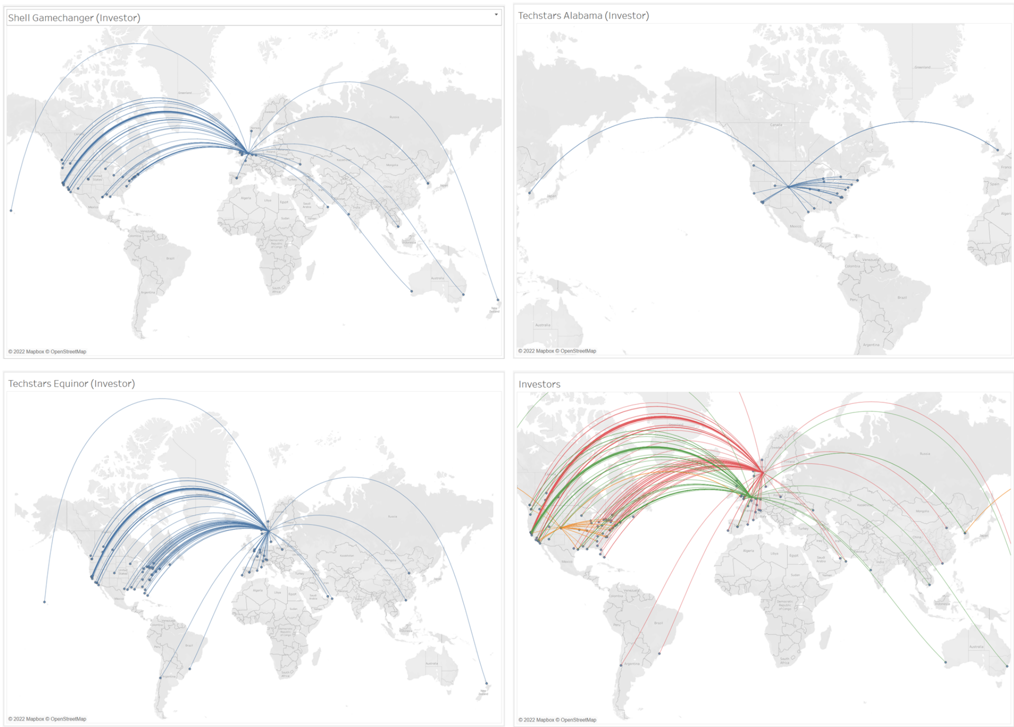
Appendix

Data matrix

Our full data matrix is available in Dropbox under this open link:

https://www.dropbox.com/sh/5m65dfxjsrdww36/AACcLEqESNjh1pD_AhRTehlJa?dl=0

Figure 1: Model A1



Aggregate	Average distance	Median	Max	Variation	Standard Deviation	Total
Investor	5,579	6,232	18,165	12,100,839	3,479	1,668,255

Shell Gamechanger	Average distance	Median	Max	Variation	Standard Deviation	Total	Network Area km2	Unique Investors	Unique Startups
Investor	**6,250**	7,707	18,165	14,026,099	3,745	**718,748**	**40,509,670**	**160**	**22**
Techstars Alabama	Average distance	Median	Max	Variation	Standard Deviation	Total	Network Area km2	Unique Investors	Unique Startups
Investor	**2,658**	2,112	9,877	5,608,622	2,368	**103,643**	**10,244,477**	**57**	**19**
Techstars Equinor	Average distance	Median	Max	Variation	Standard Deviation	Total	Network Area km2	Unique Investors	Unique Startups
Investor	**5,834**	6,463	17,210	9,721,601	3,118	**845,863**	**103,959,732**	**187**	**29**

Figure 2: Model A2

Aggregate	Total Raised	Average distance	Total
Investor	764	5,579	1,668,255

Shell Gamechanger	Total Raised	Average distance	Total	Network Area km2
Investor	**336**	6,250	**718,748**	**40,509,670**
Techstars Alabama	Total Raised	Average distance	Total	Network Area km2
Investor	**153**	2,658	**103,643**	**10,244,477**
Techstars Equinor	Total Raised	Average distance	Total	Network Area km2
Investor	**275**	5,834	**845,863**	**103,959,732**

Figure 3: Model B1

Aggregate	MeanAUM	Total AUM	Total Raised
N=257	5,248	1,348,653	764

Shell Gamechanger	MeanAUM	Total AUM	Total Raised
N= 107	**575**	61,484	**336**
Techstars Alabama	MeanAUM	Total AUM	Total Raised
N= 36	**35,307**	1,271,053	**143**
Techstars Equinior	MeanAUM	Total AUM	Total Raised
N= 114	**141**	16,116	**285**

Figure 4: Model C1

Aggregate	Average distance	Median	Max	Variation	Standard Deviation	Total
Startup	4,468	2,761	15,996	13,404,428	3,661	312,761

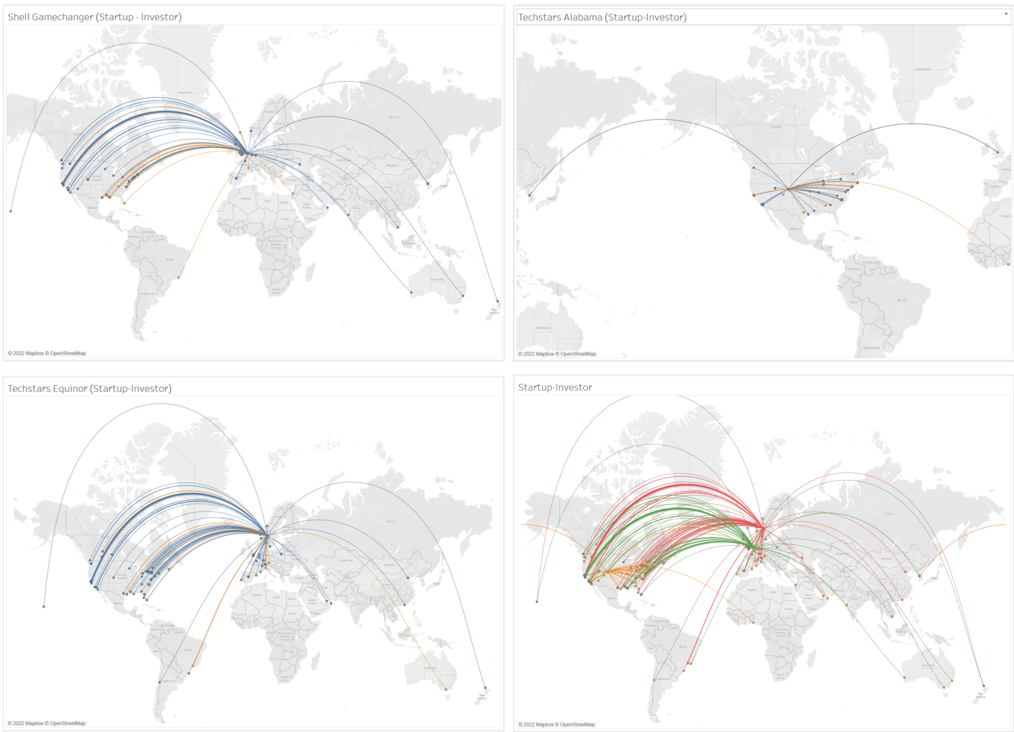
Shell Gamechanger	Average distance	Median	Max	Variation	Standard Deviation	Total	Network Area km2	Unique Investors	Unique Startups
Startup	**5,931**	7,491	9,527	10,389,715	3,223	**130,488**	**42,478,891**	**160**	**22**
Techstars Alabama	Average distance	Median	Max	Variation	Standard Deviation	Total	Network Area km2	Unique Investors	Unique Startups
Startup	**2,230**	1,792	11,106	5,313,054	2,305	**42,366**	**4,018,959**	**57**	**19**
Techstars Equinor	Average distance	Median	Max	Variation	Standard Deviation	Total	Network Area km2	Unique Investors	Unique Startups
Startup	**4,824**	5,598	15,996	16,611,285	4,076	**139,907**	**67,654,936**	**187**	**29**

Figure 5: Model C2:

Aggregate	Total Raised	Average distance	Total
Startup	764	4,468	312,761

Shell Gamechanger	Total Raised	Average distance	Total	Network Area km2
Startup	**336**	**5,931**	**130,488**	**42,478,891**
Techstars Alabama	Total Raised	Average distance	Total	Network Area km2
Startup	**153**	**2,230**	**42,366**	**4,018,959**
Techstars Equinor	Total Raised	Average distance	Total	Network Area km2
Startup	**275**	**4,824**	**139,907**	**67,654,936**

Figure 6: Model D1



	Mean	Techstars Equinor	Techstars Alabama	Shell Gamechanger
CAGR total	15 %	23 %	51 %	22 %
CAGR investors	15 %	23 %	68 %	22 %
CAGR startups	11 %	24 %	24 %	13 %
Growth total	14	43	25	7
Growth investors	12	37	19	6
Growth startups	2	6	6	1
Years	11	5	3	26
Startups final	23	29	19	22
Investors final	135	187	57	160
Total final	158	216	76	182
Startups begin	7	10	10	1
Investors begin	26	66	12	1
Total begin	33	76	22	1

Figure 7: Model E1

Shell_Gamechanger	Average Patents	Patents	Patents per Country	Countries
N=16	2.0	**32**	**1.9**	**17**
Techstars_Alabama	Average Patents	Patents	Patents per Country	Countries
N=19	0.3	**5**	**0.8**	**6**
Techstars_Equinor	Average Patents	Patents	Patents per Country	Countries
N=28	1.6	**44**	**2.0**	**22**

Figure 8: Model E2

ShellGamechanger	Average age startup	Age_startup	Countries	Average Distance	Distance
N=16	**6.9**	**110**	**3**	5,574.30	89,189.00
Techstars Alabama	Average age startup	Age_startup	Countries	Average Distance	Distance
N=19	**3.6**	**71**	**3**	2,518.80	50,375.00
Techstars Equinor	Average age startup	Age_startup	Countries	Average Distance	Distance
N=28	**5.9**	**160**	**8**	4,873.30	131,580.00

Figure 9: Model E3

Shell Gamechanger	Average age	Total age	Mean Distance	Total distance	Unique Countries
N= 107	22.4	**2,402**	6,410.4	685,916.0	**15**
Techstars Alabama	Average age	Total age	Mean Distance	Total distance	Unique Countries
N= 36	23.6	**851**	2,794.0	100,584.7	**4**
Techstars Equinior	Average age	Total age	Mean Distance	Total distance	Unique Countries
N= 114	13.4	**1,513**	6,001.7	684,196.2	**16**

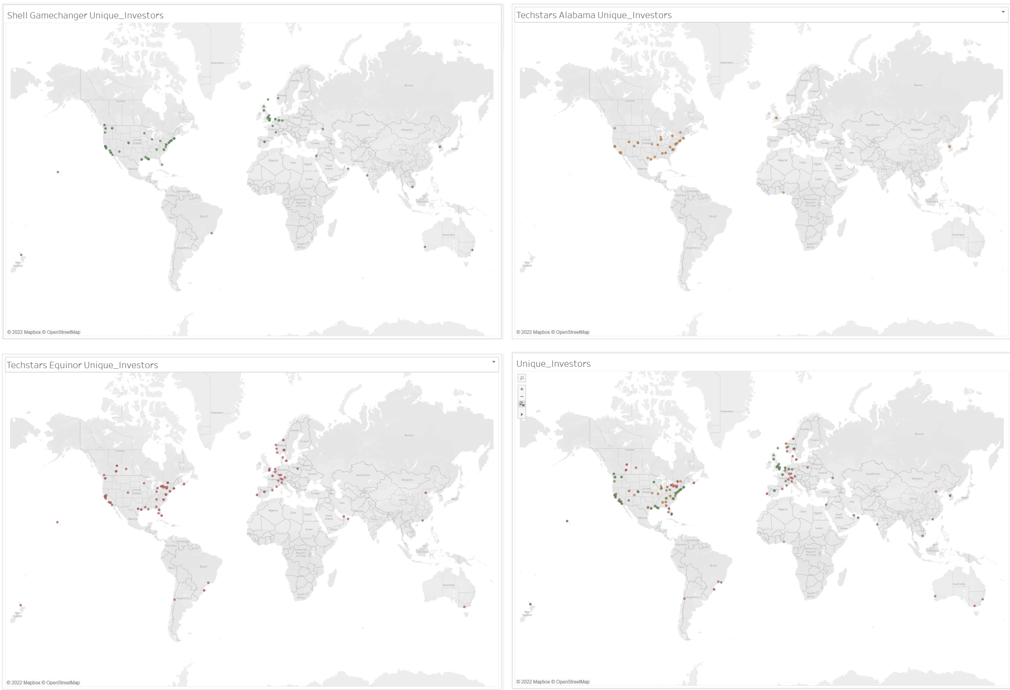
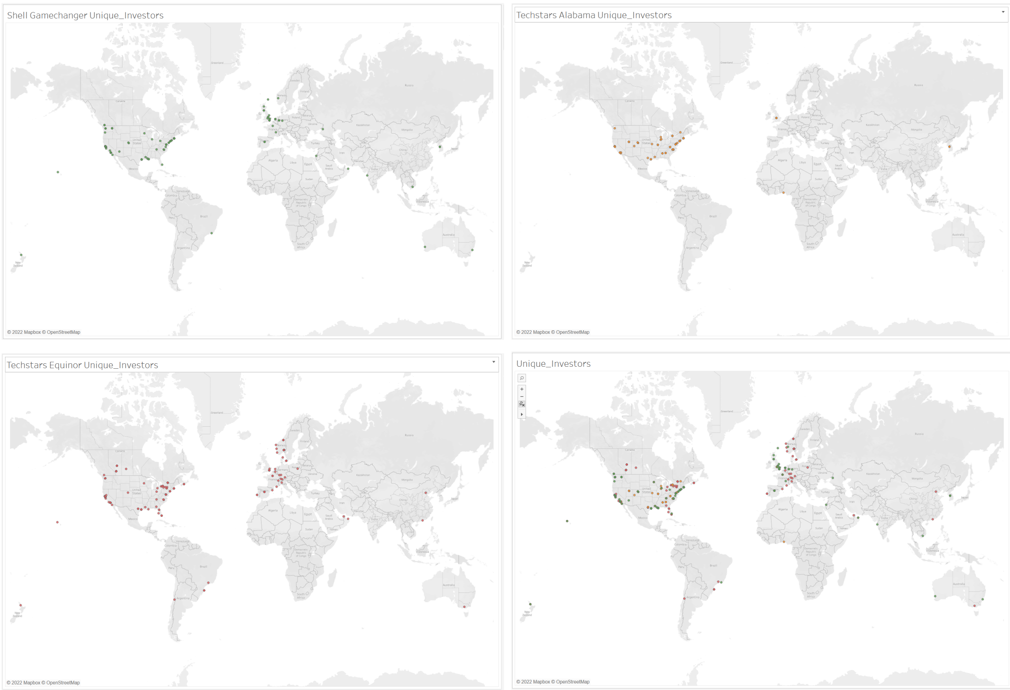


Figure 10: Model F1

Shell Gamechanger	Investors	Unique Countries	Investor/Country
	160	15	**10.7**
Techstars Alabama	Investors	Unique Countries	Investor/Country
	57	4	**14.3**
Techstars Equinior	Investors	Unique Countries	Investor/Country
	187	21	**8.9**



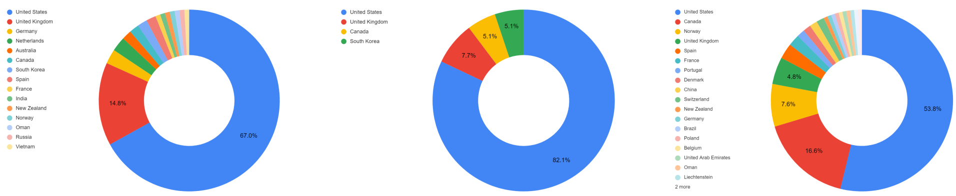


Figure 11: Model G1

	Aggregate	Average Patents	Patents	Total Raised	Avg. Raised	Avg. Raised Quartiles	Total Raised %
No patent	N=43	0	0	**234**	**5.4**	2.1	**30.6%**
Patent	N=20	4.05	81	**530**	**26.5**	3.3	**69.4%**

Shell_Gamechanger	Average Patents	Patents	Patents/Country	Total Raised	Avg. Raised	Avg. Raised Quartiles	Countries
N=16	**2.0**	**32**	1.9	**336**	**21**	2.8	17
Techstars_Alabama	Average Patents	Patents	Patents/Country	Total Raised	Avg. Raised	Avg. Raised Quartiles	Countries
N=19	**0.3**	**5**	0.8	**143**	**7.5**	2.1	6
Techstars_Equinor	Average Patents	Patents	Patents/Country	Total Raised	Avg. Raised	Avg. Raised Quartiles	Countries
N=28	**1.6**	**44**	2.0	**285**	**10.2**	2.6	22

Figure 12: Model H1

	Aggregate	Average Patents	Patents	Avg. age	Unique investors per Startup	Avg. unique investors per startup	Total Raised %
No patent	N=43	0	0	4.7	**215**	**5**	30.6%
Patent	N=20	4.05	**81**	7	**278**	**13.9**	69.4%

Shell_Gamechanger	Average Patents	Patents	Patents/Country	Avg. age	Unique Investor	Unique investors per Startup	Avg. unique investors per startup	Countries
N=16	2.0	**32**	1.9	6.9	**158**	**172**	**10.8**	17
Techstars_Alabama	Average Patents	Patents	Patents/Country	Avg. age	Unique Investor	Unique investors per Startup	Avg. unique investors per startup	Countries
N=19	0.3	**5**	0.8	3.5	**57**	**78**	**4.1**	6
Techstars_Equinor	Average Patents	Patents	Patents/Country	Avg. age	Unique Investor	Unique investors per Startup	Avg. unique investors per startup	Countries
N=28	1.6	**44**	2.0	5.9	**187**	**304**	**7.9**	22

Figure 13: Model I1

	Mean	Techstars Equinor	Techstars Alabama	Shell Gamechanger
CAGR total	15 %	23 %	51 %	22 %
CAGR investors	15 %	23 %	68 %	22 %
CAGR startups	11 %	24 %	24 %	13 %
Growth total	14	43	25	7
Growth investors	12	37	19	6
Growth startups	2	6	6	1
Years	11	5	3	26
Startups final	23	29	19	22
Investors final	135	187	57	160
Total final	158	216	76	182
Startups begin	7	10	10	1
Investors begin	26	66	12	1
Total begin	33	76	22	1

Figure 14: Model J1

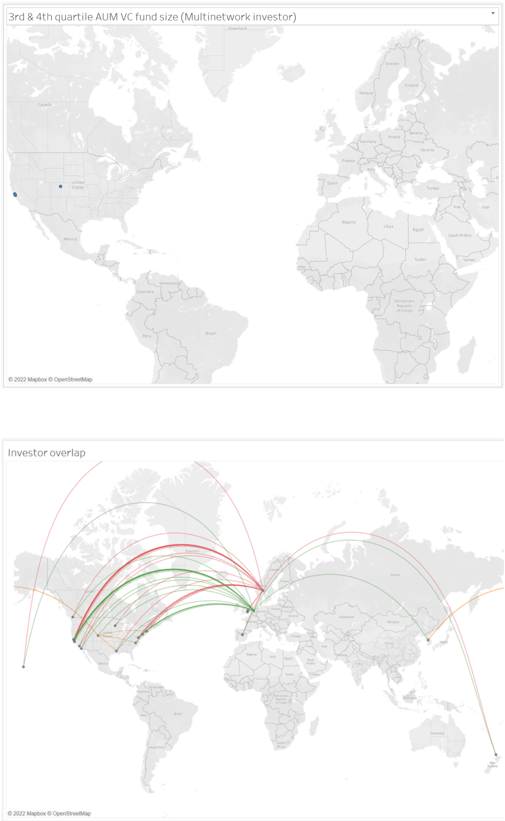


Figure 15: Model J2

Figure 16: Model K1

	Aggregate average	Average age	Total Raised	Avg. Raised	Avg. Raised Quartiles	Avg. Times accelerated	Avg. unique investors per startup	Unique investors per Startup
Below 5.4	N=33	3.3	207	6.3	2.1	1.8	5.3	175
Above 5.4	N=30	7.7	557	18.6	2.9	2.3	10.6	318

ShellGamechanger	Average age	Total Raised	Avg. Raised	Avg. Raised Quartiles	Avg. Times accelerated	Avg. unique investors per startup	Unique investors per Startup
N=16	6.875	336	21	2.75	1.875	10.75	172
Techstars Alabama	Average age	Total Raised	Avg. Raised	Avg. Raised Quartiles	Avg. Times accelerated	Avg. unique investors per startup	Unique investors per Startup
N=19	3.5	143	7.5	2.1	1.5	4.1	78
Techstars Equinor	Average age	Total Raised	Avg. Raised	Avg. Raised Quartiles	Avg. Times accelerated	Avg. unique investors per startup	Unique investors per Startup
N=28	5.9	285	10.2	2.6	2.5	8.7	243

Figure: Country profiles

SHELL GAMECHANGER

Start UP	Countries	#	%
Share	United States	15	68.18%
	United Kingdom	4	18.18%
	Brazil	1	4.55%
	France	1	4.55%
	Israel	1	4.55%
	Total	Total	Ratio
	5	22	0.23

Investors	Countries	#	%
Share	United States	77	66.96%
	United Kingdom	17	14.78%
	Germany	3	2.61%
	Netherlands	3	2.61%
	Australia	2	1.74%
	Canada	2	1.74%
	South Korea	2	1.74%
	Spain	2	1.74%
	France	1	0.87%
	India	1	0.87%
	New Zealand	1	0.87%
	Norway	1	0.87%
	Oman	1	0.87%
	Russia	1	0.87%
	Vietnam	1	0.87%
	Total	Total	Ratio
	15	115	0.13

Total Shell	Countries	#	%
Share	United States	92	67.15%
	United Kingdom	21	15.33%
	Germany	3	2.19%
	Netherlands	3	2.19%
	France	2	1.46%
	Australia	2	1.46%
	Canada	2	1.46%
	South Korea	2	1.46%
	Spain	2	1.46%
	Brazil	1	0.73%
	Israel	1	0.73%
	India	1	0.73%
	New Zealand	1	0.73%
	Norway	1	0.73%
	Oman	1	0.73%
	Russia	1	0.73%
	Vietnam	1	0.73%
	Total	Total	Ratio
	17	137	0.12

Ratio
5.23

TECHSTARS ALABAMA

Start UP	Countries	#	%
Share	United States	17	89.50%
	Nigeria	1	5.26%
	Canada	1	5.26%
	Total	Total	Ratio
	3	19	0.16

Investors	Countries	#	%
Share	United States	32	82.05%
	United Kingdom	3	7.69%
	Canada	2	5.13%
	South Korea	2	5.13%
	Total	Total	Ratio
	4	39	0.1

Total Alabama	Countries	#	%
Share	United States	49	84.48%
	Canada	3	5.17%
	United Kingdom	3	5.17%
	South Korea	2	3.45%
	Nigeria	1	1.72%
	Total	Total	Ratio
	6	58	0.1

Ratio
2.05

TECHSTARS EQUINOR

Start UP	Countries	#	%
Share	Canada	7	24.14%
	United States	6	20.69%
	Norway	5	17.24%
	United Kingdom	2	6.90%
	Germany	2	6.90%
	Brazil	2	6.90%
	Spain	2	6.90%
	Italy	1	3.45%
	Denmark	1	3.45%
	Australia	1	3.45%
	Total	Total	Ratio
	10	29	0.34

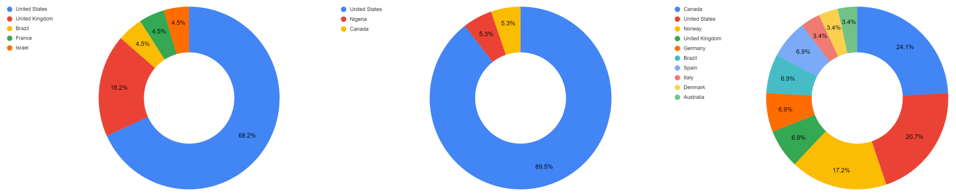
Investors	Countries	#	%
Share	United States	78	53.79%
	Canada	24	16.55%
	Norway	11	7.59%
	United Kingdom	7	4.83%
	Spain	4	2.76%
	France	3	2.07%
	Portugal	2	1.38%
	Denmark	2	1.38%
	China	2	1.38%
	Switzerland	2	1.38%
	New Zealand	1	0.69%
	Germany	1	0.69%
	Brazil	1	0.69%
	Poland	1	0.69%
	Belgium	1	0.69%
	United Arab Emirates	1	0.69%
	Oman	1	0.69%
	Liechtenstein	1	0.69%
	Netherlands	1	0.69%
	Chile	1	0.69%
	Total	Total	Ratio
	21	145	0.14

Total Equinor	Countries	#	%
Share	United States	84	48.28%
	Canada	31	17.82%
	Norway	16	9.20%
	United Kingdom	9	5.17%
	Spain	6	3.45%
	Germany	3	1.72%
	Brazil	3	1.72%
	Denmark	3	1.72%
	France	3	1.72%
	Portugal	2	1.15%
	China	2	1.15%
	Switzerland	2	1.15%
	Italy	1	0.57%
	Australia	1	0.57%
	New Zealand	1	0.57%
	Poland	1	0.57%
	Belgium	1	0.57%
	United Arab Emirates	1	0.57%
	Oman	1	0.57%
	Liechtenstein	1	0.57%
	Netherlands	1	0.57%
	Chile	1	0.57%
	Total	Total	Ratio
	22	174	0.13

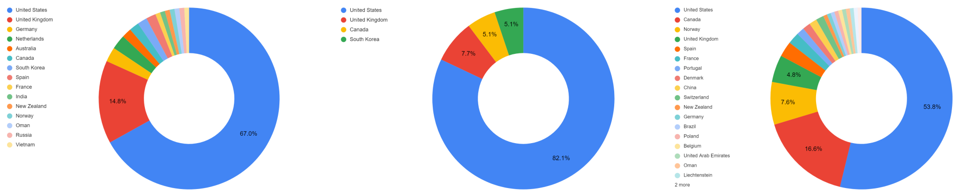
Ratio
5

Figure: Country profiles

Startup countries :



Investor Countries:



Total Countries:

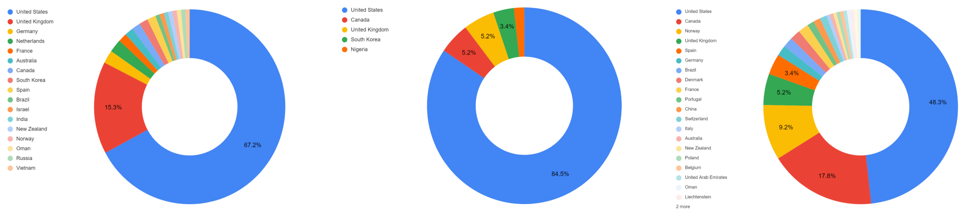
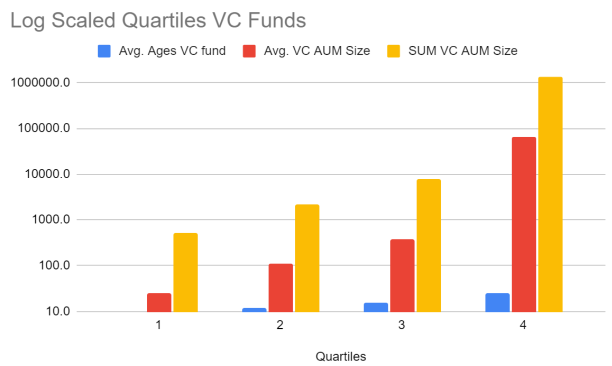


Figure: VC AUM by age/sum/quartile:

Note: Observations where VC AUM was zero was removed.



Quartiles	Avg. Ages VC fund	Avg. VC AUM Size	SUM VC AUM Size
1	9.4	24.9	522.9
2	11.8	109.8	2195.0
3	15.3	385.3	7705.0
4	25.0	66686.0	1333720.0

Distance metrics:

Techstars Alabama	Average	Median	Max	Variation	Standard Deviation	Total	Network Area km2
Startup	2,230	1,792	11,106	5,313,054	2,305	42,366	4,018,959
19	Average	Median	Max	Variation	Standard Deviation	Total	
Investor	2,658	2,112	9,877	5,608,622	2,368	103,643	10,244,477
57	Average	Median	Max	Variation	Standard Deviation	Total	
Techstars Alabama	2,517	2,045	11,106	5,457,893	2,336	146,009	

Shell Gamechanger	Average	Median	Max	Variation	Standard Deviation	Total	Network Area km2
Startup	5,931	7,491	9,527	10,389,715	3,223	130,488	42,478,891
22	Average	Median	Max	Variation	Standard Deviation	Total	
Investor	6,250	7,707	18,165	14,026,099	3,745	718,748	40,509,670
160	Average	Median	Max	Variation	Standard Deviation	Total	
ShellGamechanger	6,199	7,702	18,165	13,375,258	3,657	849,236	

Equinor	Average	Median	Max	Variation	Standard Deviation	Total	Network Area km2
Startup	4,824	5,598	15,996	16,611,285	4,076	139,907	67,654,936
28	Average	Median	Max	Variation	Standard Deviation	Total	
Investor	5,834	6,463	17,210	9,721,601	3,118	845,863	103,959,732
187	Average	Median	Max	Variation	Standard Deviation	Total	
Equinor	5,665	6,450	17,210	10,922,765	3,305	985,770	

Aggregate	Average	Median	Max	Variation	Standard Deviation	Total
Startup	4,468	2,761	15,996	13,404,428	3,661	312,761
	Average	Median	Max	Variation	Standard Deviation	Total
Investor	5,579	6,232	18,165	12,100,839	3,479	1,668,255
	Average	Median	Max	Variation	Standard Deviation	Total
All	5,369	6,232	18,165	12,502,776	3,536	1,981,016

Distance between points in meters = ACOS(SIN(LAT1*PI()/180)*SIN(LAT2*PI()/180) + COS(LAT1*PI()/180)*COS(LAT2*PI()/180)*COS(LON2*PI()/180-LON1*PI()/180)) * 6371000

Figure: Pearson's R correlation matrix:

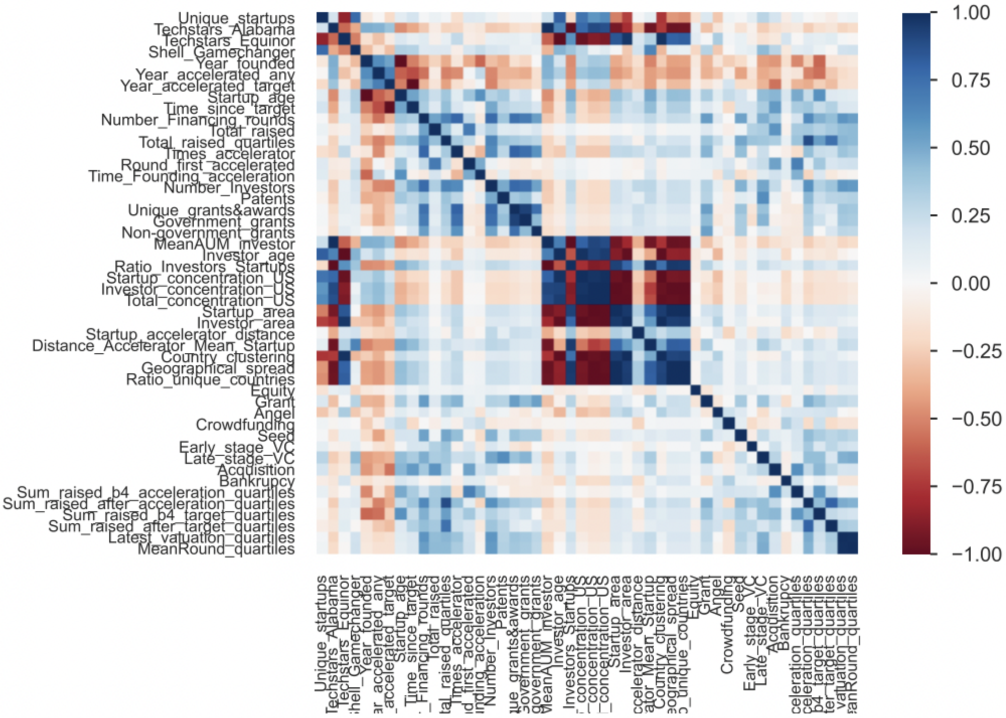
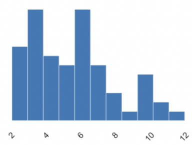


Figure: Descriptions of chosen variables generated in Python:

- **Age startups**

Age_startup
Real number ($\mathbb{R}_{>0}$)
HIGH CORRELATION

Distinct	11	Mean	5.412698413
Distinct (%)	17.5%	Minimum	2
Missing	0	Maximum	12
Missing (%)	0.0%	Zeros	0
Infinite	0	Zeros (%)	0.0%
Infinite (%)	0.0%	Memory size	504.0 B

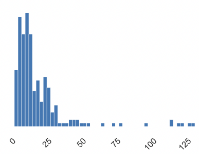


Toggle details

- **Age investors**

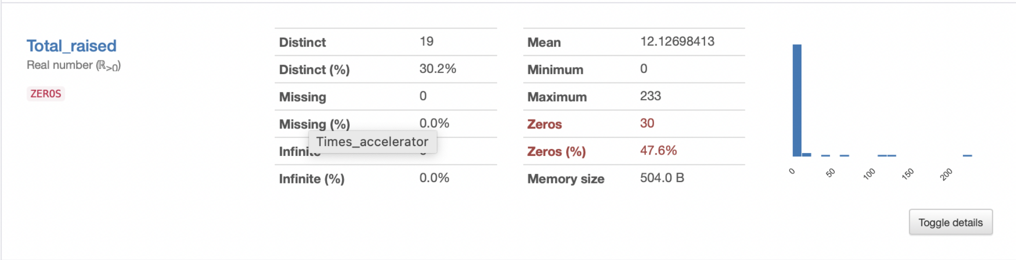
Age_VC
Real number ($\mathbb{R}_{\geq 0}$)
MISSING

Distinct	51	Mean	18.64220183
Distinct (%)	23.4%	Minimum	1
Missing	214	Maximum	132
Missing (%)	49.5%	Zeros	0
Infinite	0	Zeros (%)	0.0%
Infinite (%)	0.0%	Memory size	3.4 KiB



Toggle details

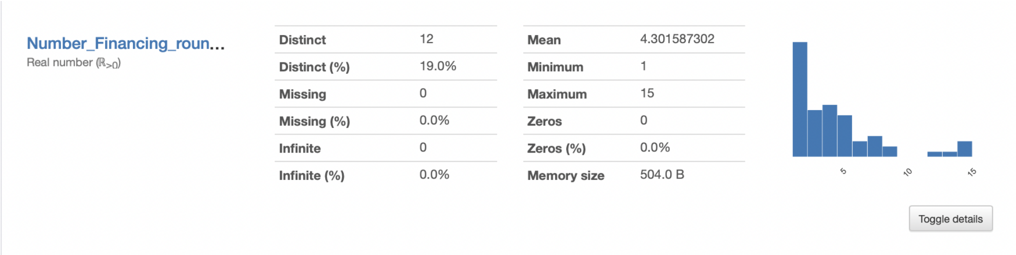
- **Capital raised**



Distinct	19
Distinct (%)	30.2%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%

Mean	12.12698413
Minimum	0
Maximum	233
Zeros	30
Zeros (%)	47.6%
Memory size	504.0 B

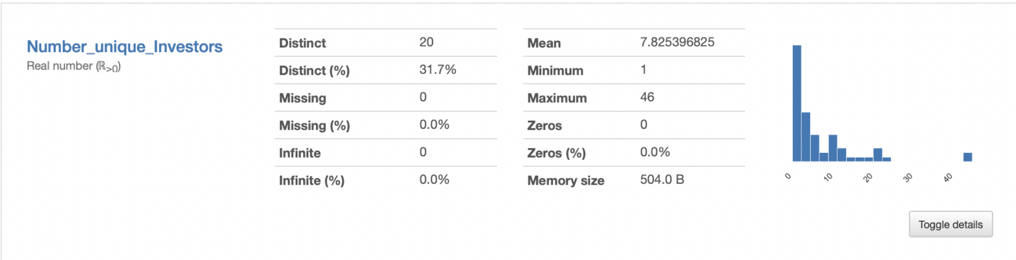
- **Financing rounds**



Distinct	12
Distinct (%)	19.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%

Mean	4.301587302
Minimum	1
Maximum	15
Zeros	0
Zeros (%)	0.0%
Memory size	504.0 B

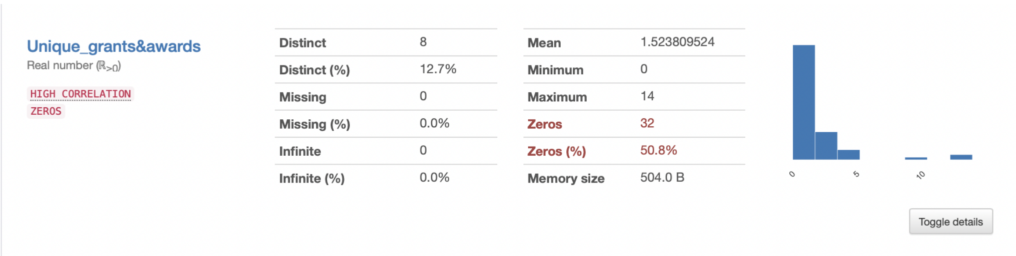
- **Total investors**



Distinct	20
Distinct (%)	31.7%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%

Mean	7.825396825
Minimum	1
Maximum	46
Zeros	0
Zeros (%)	0.0%
Memory size	504.0 B

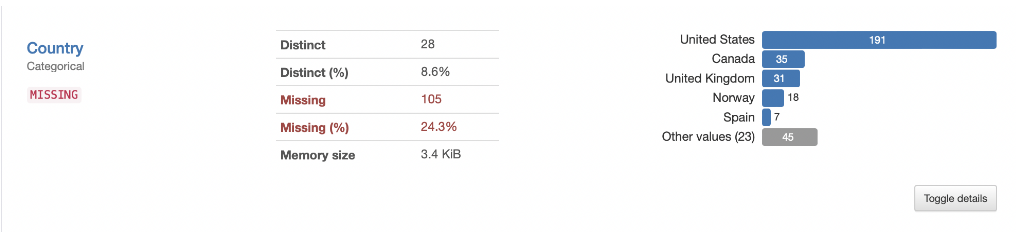
- **Grants and awards**



Distinct	8
Distinct (%)	12.7%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%

Mean	1.523809524
Minimum	0
Maximum	14
Zeros	32
Zeros (%)	50.8%
Memory size	504.0 B

- **Country distribution**

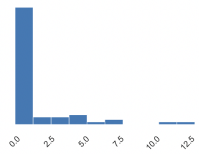


Distinct	28
Distinct (%)	8.6%
Missing	105
Missing (%)	24.3%
Memory size	3.4 KiB

- **Patent distribution**

Patents
Real number ($\mathbb{R}_{\geq 0}$)
ZEROS

Distinct	10	Mean	1.285714286
Distinct (%)	15.9%	Minimum	0
Missing	0	Maximum	13
Missing (%)	0.0%	Zeros	43
Infinite	0	Zeros (%)	68.3%
Infinite (%)	0.0%	Memory size	504.0 B

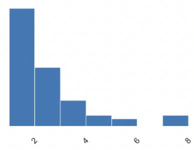


Toggle details

- **Times accelerated**

Times_accelerator
Real number ($\mathbb{R}_{\geq 0}$)

Distinct	7	Mean	2.047619048
Distinct (%)	11.1%	Minimum	1
Missing	0	Maximum	8
Missing (%)	0.0%	Zeros	0
Infinite	0	Zeros (%)	0.0%
Infinite (%)	0.0%	Memory size	504.0 B



Toggle details

Tables: Excerpt of initial regression outputs

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     Country_clustering   R-squared:                       1.000
Model:                            OLS   Adj. R-squared:                  1.000
Method:                 Least Squares   F-statistic:                 4.423e+30
Date:                Wed, 08 Jun 2022   Prob (F-statistic):               0.00
Time:                        15:38:13   Log-Likelihood:                 1984.1
No. Observations:                  63   AIC:                            -3958.
Df Residuals:                      58   BIC:                            -3948.
Df Model:                           4
Covariance Type:                  HAC
===============================================================================================
                               coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------------------
Intercept                    5.9204     6.8e-15    8.7e+14      0.000       5.920       5.920
Ratio_Investors_Startups     0.3060    6.43e-16   4.76e+14      0.000       0.306       0.306
Startup_concentration_US    -1.2518     1.4e-15  -8.92e+14      0.000      -1.252      -1.252
Investor_concentration_US    3.4045    3.94e-15   8.64e+14      0.000       3.404       3.404
Ratio_unique_countries       0.6711    7.67e-16   8.75e+14      0.000       0.671       0.671
MeanAge_VC                  -0.3164    3.42e-16  -9.24e+14      0.000      -0.316      -0.316
Government_grants        -1.388e-15    7.02e-16     -1.976      0.053   -2.79e-15     1.8e-17
Patents                   2.914e-16    3.46e-16      0.841      0.404   -4.02e-16    9.85e-16
==============================================================================
Omnibus:                       25.855   Durbin-Watson:                   0.533
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               52.128
Skew:                           1.324   Prob(JB):                     4.79e-12
Kurtosis:                       6.584   Cond. No.                     1.76e+18
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 8 lags and without small sample
correction
[2] The smallest eigenvalue is 7.63e-33. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     Geographical_spread   R-squared:                       1.000
Model:                             OLS   Adj. R-squared:                  1.000
Method:                  Least Squares   F-statistic:                 8.686e+16
Date:                 Wed, 08 Jun 2022   Prob (F-statistic):               0.00
Time:                         15:38:23   Log-Likelihood:                 1030.9
No. Observations:                   63   AIC:                            -2054.
Df Residuals:                       59   BIC:                            -2045.
Df Model:                            3
Covariance Type:                   HAC
===================================================================================================================
                                                    coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------------------------------------
Intercept                                         4.9444    1.13e-08   4.37e+08      0.000       4.944       4.944
Distance_Accelerator_Mean_Investor[T.5 834]      -0.8024    1.83e-09  -4.37e+08      0.000      -0.802      -0.802
Distance_Accelerator_Mean_Investor[T.6 250]       0.8110    1.85e-09   4.37e+08      0.000       0.811       0.811
Accelerator_startup_Area                       2.659e-07    1.04e-15   2.57e+08      0.000    2.66e-07    2.66e-07
Accelerator_investor_Area                     -1.292e-09    6.82e-16   -1.9e+06      0.000   -1.29e-09   -1.29e-09
Distance_Accelerator_Startup                   1.334e-17    5.88e-13   2.27e-05      1.000   -1.18e-12    1.18e-12
==============================================================================
Omnibus:                       13.378   Durbin-Watson:                   0.003
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               10.256
Skew:                           0.870   Prob(JB):                      0.00593
Kurtosis:                       2.062   Cond. No.                     5.52e+23
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 8 lags and without small sample
correction
[2] The smallest eigenvalue is 1.59e-30. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     Latest_valuation_quartiles   R-squared:                       0.444
Model:                                    OLS   Adj. R-squared:                  0.373
Method:                         Least Squares   F-statistic:                     33.52
Date:                        Wed, 08 Jun 2022   Prob (F-statistic):           1.21e-17
Time:                                15:38:31   Log-Likelihood:                -92.761
No. Observations:                          63   AIC:                             201.5
Df Residuals:                              55   BIC:                             218.7
Df Model:                                   7
Covariance Type:                          HAC
=====================================================================================================
                                        coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------------------------
Intercept                            -0.1004      0.191     -0.525      0.602      -0.484       0.283
C(Seed)[T.1]                          0.1190      0.362      0.329      0.744      -0.606       0.844
C(Early_stage_VC)[T.1]                0.3653      0.367      0.997      0.323      -0.369       1.100
C(Late_stage_VC)[T.1]                 1.1811      0.544      2.173      0.034       0.092       2.271
Number_Financing_rounds               0.1299      0.078      1.669      0.101      -0.026       0.286
Number_unique_Investors               0.0244      0.017      1.435      0.157      -0.010       0.059
Sum_raised_after_target_quartiles     0.1032      0.086      1.197      0.236      -0.070       0.276
Sum_raised_b4_target_quartiles       -0.1566      0.099     -1.585      0.119      -0.354       0.041
==============================================================================
Omnibus:                        0.877   Durbin-Watson:                   1.830
Prob(Omnibus):                  0.645   Jarque-Bera (JB):                0.902
Skew:                           0.265   Prob(JB):                        0.637
Kurtosis:                       2.750   Cond. No.                         50.2
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 8 lags and without small sample
correction
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:        Network_Quality   R-squared:                       1.000
Model:                            OLS   Adj. R-squared:                  1.000
Method:                 Least Squares   F-statistic:                 5.489e+28
Date:                Wed, 08 Jun 2022   Prob (F-statistic):               0.00
Time:                        15:56:28   Log-Likelihood:                 1551.1
No. Observations:                  63   AIC:                            -3090.
Df Residuals:                      57   BIC:                            -3077.
Df Model:                           5
Covariance Type:                  HAC
===========================================================================================
                              coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------------
Intercept                   0.0441   8.04e-16   5.48e+13      0.000       0.044       0.044
Age_startup              6.162e-15   1.24e-13      0.049      0.961   -2.43e-13    2.55e-13
MeanAge_VC                  0.3536   3.17e-14   1.12e+13      0.000       0.354       0.354
Country_clustering          0.2929   9.93e-15   2.95e+13      0.000       0.293       0.293
Geographical_spread         1.0848   2.18e-14   4.98e+13      0.000       1.085       1.085
Number_unique_Investors -2.637e-16   2.65e-14     -0.010      0.992   -5.33e-14    5.28e-14
MeanAUM_VC               4.577e-05   9.06e-17   5.05e+11      0.000    4.58e-05    4.58e-05
Total_raised_quartiles  -1.388e-15   3.42e-13     -0.004      0.997   -6.87e-13    6.84e-13
==============================================================================
Omnibus:                       28.483   Durbin-Watson:                   0.099
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               11.909
Skew:                           0.861   Prob(JB):                      0.00259
Kurtosis:                       1.747   Cond. No.                     2.21e+20
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 8 lags and without small sample
correction
[2] The smallest eigenvalue is 1.69e-30. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
```

Text: Python code script

We are making our code available, though not our preliminary figures and regression outputs. Since we have run the code in Jupyter, copy-pasting the text format presented here will probably run.

```
#This Python script has been run in Jupyter, through the Anaconda Navigator

----
#Import all libraries and packages and set as variables

import numpy as np

import pandas as pd
from pandas import Series, DataFrame
import pandas_datareader as pdr
from pandas_profiling import ProfileReport

import matplotlib.pyplot as plt
%matplotlib inline

from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures

import statsmodels as sm
import statsmodels.formula.api as smf
import statsmodels.stats.api as sms

import scipy.stats as stats

----

#Read in the data file
#We import the CSV file from the same folder as our script file is located.

df = pd.read_csv('Finaldataset.csv', index_col=None, na_values=['NA'])
----

#View the data file
df

#View the header

df.head()

#View the shape

df.shape

#All looked as it was suppose to

----
#Clean the data matrix if needed. We did that when building it in Excel so less was needed to perform i Python.
#We did need to rename a variable since it contained a 'space' which caused reading troubles for Python

#Rename variable
df.rename(columns = {'TechStars_Alabama ':'TechStars_Alabama'}, inplace = True)

----
#We did more data exploration
#Pandas ProfileReport outputs a lot of info on the data set
#and is very useful to uncover inconsistencies in the data
#The correlation matrix and variable frequency tables and histograms comes from this code
```

```
profile = ProfileReport(df, title="Pandas Profiling Report")
profile.to_widgets()
profile.to_notebook_iframe()

#We then saved the generated file

profile.to_file("Finaldatasetreport")

----
#We used Matplotlib to plot out variables and play with the data to see how it interacted
#Please understand that we have only included general commands, not all of the commands we used.

#Typically, it looked like this, with one or several variables
plt.figure(figsize=(6, 4))
plt.plot(df.Observation, df.Observations, "ro", label= 'Observations') #first variable
plt.plot(df.Observation, df.Age_startup, "go", label= 'Age') #second variable

plt.title("This is your title")
plt.xlabel('Name1')
plt.ylabel('Name2')
plt.show()

#We did similar plots too
plt.bar() #barplot
plt.hist() #histogram
plt.boxplot() #boxplot
----

#Then we performed simple linear regressions with OLS
#to uncover relationships in the data
#Still part of the data exploration phase

#Typically, it looked like this

linreg = LinearRegression()

#Set the values
x = df['Observation'].values
y = df['Age_startup'].values

#Restructure the X values
x = x.reshape(-1, 1)

#Regression output
linreg.fit(x,y)

#Create new variable for output
y_pred = linreg.predict(x)

#Plot the regression
plt.scatter(x,y, color='k')
plt.plot(x, y_pred, color='r')
plt.title("This is your title")
plt.xlabel('Name1')
plt.ylabel('Name2')
plt.show()

#Print numerical values of the function Y=B0+B1X1+u
print(linreg.coef_)
print(linreg.intercept_)

----
#We experimented with different functional forms by changing the degree

#Typically it looked like this

#Set the degrees of freedom to fi the line to the data better
poly = PolynomialFeatures(degree=2)
x_poly = poly.fit_transform(x)
```

```
#Regression output
linreg.fit(x_poly, y)

#Create new variable for output
y_pred = linreg.predict(x_poly)

#Plot the regression
plt.scatter(x,y, color='blue')
plt.plot(x, y_pred, color='red')
plt.title("This is your title")
plt.xlabel('Name1')
plt.ylabel('Name2')
plt.show()


print(linreg.coef_)
print(linreg.intercept_)
----


#When we were through with data exploration and cleaning, we
#headed for our regression tables.

#We used OLS multiple variable regression from Statsmodels package.
#Controlled standard errors by using HAC to make them heteroscedasticity and autocorrelation robust.
#It improved the results to add HAC.

#Here we will add all our code used. A total of 16 operations.

----

#Network quality, Full sample

mod = smf.ols(formula="Network_Quality ~ Age_startup + MeanAge_VC + Country_clustering +
Geographical_spread + Number_unique_Investors + MeanAUM_VC + Total_raised_quartiles",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
----

#Network quality, TechStars Equinor

mod = smf.ols(formula="Network_Quality ~ C(TechStars_Equinor) + Age_startup + MeanAge_VC +
Country_clustering + Geographical_spread + Number_unique_Investors + MeanAUM_VC + Total_raised_quartiles",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
----

#Network quality, TechStars Alabama

mod = smf.ols(formula="Network_Quality ~ C(TechStars_Alabama) + Age_startup + MeanAge_VC +
Country_clustering + Geographical_spread + Number_unique_Investors + MeanAUM_VC + Total_raised_quartiles",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
----

#Network quality, Shell Gamechanger

mod = smf.ols(formula="Network_Quality ~ C(Shell_Gamechanger) + Age_startup + MeanAge_VC +
Country_clustering + Geographical_spread + Number_unique_Investors + MeanAUM_VC + Total_raised_quartiles",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
----

#Country clustering, Full sample

mod = smf.ols(formula="Country_clustering ~ Ratio_Investors_Startups + Startup_concentration_US +
Investor_concentration_US + Ratio_unique_countries + MeanAge_VC + Government_grants + Patents",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Country clustering, TechStars Equinor

```
mod = smf.ols(formula="Country_clustering ~ C(TechStars_Equinor) + Ratio_Investors_Startups +
Startup_concentration_US + Investor_concentration_US + Ratio_unique_countries + MeanAge_VC +
Government_grants + Patents", data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Country clustering, TechStars Alabama

```
mod = smf.ols(formula="Country_clustering ~ C(TechStars_Alabama) + Ratio_Investors_Startups +
Startup_concentration_US + Investor_concentration_US + Ratio_unique_countries + MeanAge_VC +
Government_grants + Patents", data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Country clustering, Shell Gamechanger

```
mod = smf.ols(formula="Country_clustering ~ C(Shell_Gamechanger) + Ratio_Investors_Startups +
Startup_concentration_US + Investor_concentration_US + Ratio_unique_countries + MeanAge_VC +
Government_grants + Patents", data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Geographical spread, Full sample

```
mod = smf.ols(formula="Geographical_spread ~ Accelerator_startup_Area + Accelerator_investor_Area +
Distance_Accelerator_Startup + Distance_Accelerator_Mean_Investor",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Geographical spread, TechStars Equinor

```
mod = smf.ols(formula="Geographical_spread ~ C(TechStars_Equinor) + Accelerator_startup_Area +
Accelerator_investor_Area + Distance_Accelerator_Startup + Distance_Accelerator_Mean_Investor",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Geographical spread, TechStars Alabama

```
mod = smf.ols(formula="Geographical_spread ~ C(TechStars_Alabama) + Accelerator_startup_Area +
Accelerator_investor_Area + Distance_Accelerator_Startup + Distance_Accelerator_Mean_Investor",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Geographical spread, Shell Gamechanger

```
mod = smf.ols(formula="Geographical_spread ~ C(Shell_Gamechanger) + Accelerator_startup_Area +
Accelerator_investor_Area + Distance_Accelerator_Startup + Distance_Accelerator_Mean_Investor",
data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Latest valuation, Full sample

```
mod = smf.ols(formula="Latest_valuation_quartiles ~ C(Seed) + C(Early_stage_VC) + C(Late_stage_VC) +
Number_Financing_rounds + Number_unique_Investors + Sum_raised_after_target_quartiles +
Sum_raised_b4_target_quartiles", data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
```

#Latest valuation, TechStars Equinor

```
mod = smf.ols(formula="Latest_valuation_quartiles ~ C(TechStars_Equinor) + C(Seed) + C(Early_stage_VC) +
C(Late_stage_VC) + Number_Financing_rounds + Number_unique_Investors + Sum_raised_after_target_quartiles +
Sum_raised_b4_target_quartiles", data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
```

```
print(mod.summary())
----

#Latest valuation, TechStars Alabama

mod = smf.ols(formula="Latest_valuation_quartiles ~ C(TechStars_Alabama) + C(Seed) + C(Early_stage_VC) +
C(Late_stage_VC) + Number_Financing_rounds + Number_unique_Investors + Sum_raised_after_target_quartiles +
Sum_raised_b4_target_quartiles", data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
----

#Latest valuation, Shell Gamechanger

mod = smf.ols(formula="Latest_valuation_quartiles ~ C(Shell_Gamechanger) + C(Seed) + C(Early_stage_VC) +
C(Late_stage_VC) + Number_Financing_rounds + Number_unique_Investors + Sum_raised_after_target_quartiles +
Sum_raised_b4_target_quartiles", data=df).fit().get_robustcov_results(cov_type='HAC', maxlags=8)
print(mod.summary())
----
#End of code script :)
----
```

Text: Description of Python packages:

Numpy.org

- Numpy is a fundamental package for scientific computing with Python
- NumPy can be used to perform a wide variety of mathematical operations on arrays. It adds powerful data structures to Python that guarantee efficient calculations with arrays and matrices and it supplies an enormous library of high-level mathematical functions that operate on these arrays and matrices.
- Citation:
    - Oliphant, T. E. (2015). Guide to NumPy. Continuum Press.

Pandas.pydata.org

- Pandas is a fast, powerful, flexible, and easy-to-use open-source data analysis and manipulation tool, built on top of the Python programming language.
- It is most widely used for data science and data analysis tasks
- From the pandas library we imported Series, Dataframe, and Datareader. All packages to help with the basic data analysis work.
- From Pandas_profiling we imported ProfileReport, an advanced packaged that allowes the user to quickly extract key figures from a data set in the form of a report.
- Citations:
    - Jeff Reback, Wes McKinney, jbrockmendel, Joris Van den Bossche, Tom Augspurger, Phillip Cloud, gfyoung, Sinhrks, Adam Klein, Matthew Roeschke, Simon Hawkins, Jeff Tratner, Chang She, William Ayd, Terji Petersen, Marc Garcia, Jeremy Schendel, Andy Hayden, MomIsBestFriend, … Mortada Mehyar. (2020). pandas-dev/pandas: Pandas 1.0.3 (v1.0.3). Zenodo. https://doi.org/10.5281/zenodo.3715232
    - McKinney, Proceedings of the 9th Python in Science Conference, Volume 445, 2010.

Matplotlib.org

- Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib allows the user to graph and present the data in numerous ways.
- The package pyplot is a plotting library used for 2D graphics, useful for plotting, histograms, linear regression visualizations, and more.
- Citation:
    - J. D. Hunter, "Matplotlib: A 2D Graphics Environment", Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, 2007.

Scikit-learn.org

- Sklearn is a simple and efficient tool for predictive data analysis
- We imported the packages LinearRegression and PolynomialFeatures to perform linear regression and polynomial regressions. Then we used Matplotlib to graphically plot them out.
- Citation:
    - Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Statsmodels.org

- Statsmodels is a Python module that provides classes and functions for the estimation of many different statistical models, as well as for conducting statistical tests, and statistical data exploration.
- We imported packages to help us perform linear OLS regressions on our data sets and then print the finished regression outputs.
- Citation:
    - Seabold, Skipper, and Josef Perktold. “statsmodels: Econometric and statistical modeling with python.” Proceedings of the 9th Python in Science Conference. 2010.

Literature list

Ahuja, G. (2000). *Collaboration Networks, Structural Holes, and Innovation:A Longitudinal Study*. 31.

Aldrich, H., & Ruef, M. (2006). *Organizations Evolving*. https://doi.org/10.4135/9781446212509

Altman, N., & Krzywinski, M. (2015). Association, correlation and causation. *Nature Methods*, *12*(10), 899–900. https://doi.org/10.1038/nmeth.3587

Barnett, W. P., & Hansen, M. T. (1996). The red queen in organizational evolution. *Strategic Management Journal*, *17*(S1), 139–157.

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Basu, S., Phelps, C., & Kotha, S. (2011). Towards understanding who makes corporate venture capital investments and why. *Journal of Business Venturing*, *26*(2), 153–171.

Baum, J. A. C., Calabrese, T., & Silverman, B. S. (2000). *Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology*. 28.

Bell, G. G. (2005). Clusters, networks, and firm innovativeness. *Strategic Management Journal*, *26*(3), 287–295.

Benson, D., & Ziedonis, R. H. (2009). Corporate venture capital as a window on new technologies: Implications for the performance of corporate investors when acquiring startups. *Organization Science*, *20*(2), 329–351.

Bernstein, S., Giroud, X., & Townsend, R. R. (2016). The impact of venture capital monitoring. *The Journal of Finance*, *71*(4), 1591–1622.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network Analysis in the Social Sciences. *Science*, *323*(5916), 892–895. https://doi.org/10.1126/science.1165821

Breiger, R. (2004). *The analysis of social networks*. https://scholar.google.com/scholar_lookup?title=The%20analysis%20of%20social%20networks&author=RL.%20Breiger&pages=505-526&publication_year=2004

Bruderl, J., & Schussler, R. (1990). Organizational Mortality: The Liabilities of Newness and Adolescence. *Administrative Science Quarterly*, *35*(3), 530. https://doi.org/10.2307/2393316

Campbell, B. (1978). Richard E. Leakey and Roger Lewin: Origins: What new discoveries reveal about the emergence of our species and its possible future. London: Macdonald and Jane's, 1977. 264 pp., numerous unnumbered illustrations. £8.95. *Antiquity*, *52*(206), 249–250. https://doi.org/10.1017/S0003598X00072318

Capaldo, A. (2007). Network structure and innovation: The leveraging of a dual network as a distinctive relational capability. *Strategic Management Journal*, *28*(6), 585–608.

Carlile, P. R. (2004). Transferring, Translating, and Transforming: An Integrative Framework for Managing Knowledge Across Boundaries. *Organization*

Science, *15*(5), 555–568. https://doi.org/10.1287/orsc.1040.0094

Carpenter, M. A., Li, M., & Jiang, H. (2012). Social Network Research in Organizational Contexts: A Systematic Review of Methodological Issues and Choices. *Journal of Management*, *38*(4), 1328–1361. https://doi.org/10.1177/0149206312440119

Carrington, P., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis* (Vol. 28). Cambridge university press. https://scholar.google.com/scholar_lookup?hl=en&publication_year=2005&author=PJ+Carrington&author=J+Scott&author=S+Wasserman&title=Models+and+Methods+in+Social+Network+Analysis

CARROLL, G. R., & SWAMINATHAN, A. (1992). The Organizational Ecology of Strategic Groups in the American Brewing Industry from 1975 to 1990 *. *Industrial and Corporate Change*, *1*(1), 65–97. https://doi.org/10.1093/icc/1.1.65

Casciaro, T., & Piskorski, M. J. (2005). Power Imbalance, Mutual Dependence, and Constraint Absorption: A Closer Look at Resource Dependence Theory. *Administrative Science Quarterly*, *50*(2), 167–199. https://doi.org/10.2189/asqu.2005.50.2.167

Cassar, G. (2004). The financing of business start-ups. *Journal of Business Venturing*, *19*(2), 261–283. https://doi.org/10.1016/S0883-9026(03)00029-6

Chemmanur, T. J., Loutskina, E., & Tian, X. (2014). Corporate venture capital, value creation, and innovation. *The Review of Financial Studies*, *27*(8), 2434–2473.

Chesbrough, H. (2002). Making Sense of Corporate Venture Capital. *Harvard Business Review*. https://hbr.org/2002/03/making-sense-of-corporate-venture-capital

Coase, R. H. (1937). The Nature of the Firm. *Economica*, *4*(16), 386–405. https://doi.org/10.1111/j.1468-0335.1937.tb00002.x

Cohen, D. G., & Feld, B. (2011). *Do more faster: TechStars lessons to accelerate your startup*. Wiley.

Crossley, N. (2010). The social world of the network. Combining qualitative and quantitative elements in social network analysis. *Sociologica*. https://scholar.google.com/scholar_lookup?hl=en&publication_year=2010&author=N+Crossley&title=The+social+world+of+the+network%3A+Combining+qualitative+and+quantitative+elements+in+social+network+analysis&doi=10.2383%2F32049

Currie, R. A. F., Ault, G. W., Foote, C. E. T., Burt, G. M., & McDonald, J. R. (2004). Fundamental research challenges for active management of distribution networks with high levels of renewable generation. *39th International Universities Power Engineering Conference, 2004. UPEC 2004.*, *3*, 1024–1028.

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm* (Vol. 2). Englewood Cliffs, NJ.

Darwin, C. (2003). *The origin of species: By means of natural selection of The preservation of favored races in the struggle for life* (J. Huxley, Ed.; 150th anniversary editien). Signet classics.

DeGennaro, R. P., & Dwyer, G. P. (2014). Expected returns to stock investments by angel investors in groups. *European Financial Management*, *20*(4), 739–755.

Dushnitsky, G., & Lenox, M. J. (2005). When do firms undertake R&D by investing in new ventures? *Strategic Management Journal*, *26*(10), 947–965. https://doi.org/10.1002/smj.488

Dushnitsky, G., & Shapira, Z. (2010). Entrepreneurial finance meets organizational reality: Comparing investment practices and performance of corporate and independent venture capitalists. *Strategic Management Journal*, *31*(9), 990–1017.

Dyer, J. H., & Singh, H. (1998). The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage. *Academy of Management Review*, *23*(4), 660–679. https://doi.org/10.5465/amr.1998.1255632

Erickson, B. H., & Nosanchuk, T. A. (1983). Applied network sampling. *Social Networks*, *5*(4), 367–382.

Feld, B., & Mendelson, J. (2019). *Venture deals: Be smarter than your lawyer and venture capitalist* (Fourth edition). John Wiley & Sons, Inc.

Gnyawali, D. R., & Park, B.-J. R. (2011). Co-opetition between giants: Collaboration with competitors for technological innovation. *Research Policy*, *40*(5), 650–663.

Gompers, P., Kovner, A., Lerner, J., & Scharfstein, D. (2010). Performance persistence in entrepreneurship. *Journal of Financial Economics*, *96*(1), 18–32.

Granovetter, M. (1976). Network sampling: Some first steps. *American Journal* *of Sociology*, *81*(6), 1287–1303.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122.

Gulati, R., Lavie, D., & Singh, H. (2009). The nature of partnering experience and the gains from alliances. *Strategic Management Journal*, *30*(11), 1213–1233. https://doi.org/10.1002/smj.786

Gulati, R., Nohria, N., & Zaheer, A. (2000). Strategic networks. *Strategic Management Journal*, *21*(3), 203–215. https://doi.org/10.1002/(SICI)1097-0266(200003)21:3<203::AID-SMJ102>3.0.CO;2-K

Hallen, B., Bingham, C., & Cohen, S. (2014). Do Accelerators Accelerate? A Study of Venture Accelerators as a Path to Success? *Academy of Management Proceedings*, *2014*, 12955–12955. https://doi.org/10.5465/AMBPP.2014.185

Hallen, B. L. (2008). The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments? *Administrative Science Quarterly*, *53*(4), 685–718.

Hannan, M. T., & Freeman, J. (1977). The population ecology of organizations. *American Journal of Sociology*, *82*(5), 929–964.

Helfat, C. E., & Raubitschek, R. S. (2018). Dynamic and integrative capabilities for profiting from innovation in digital platform-based ecosystems. *Research Policy*, *47*(8), 1391–1399.

Hillman, A. J., Withers, M. C., & Collins, B. J. (2009). Resource Dependence Theory: A Review. *Journal of Management*, *35*(6), 1404–1427.

https://doi.org/10.1177/0149206309343469

Hochberg, Y. V., Ljungqvist, A., & Lu, Y. (2007). Whom you know matters: Venture capital networks and investment performance. *The Journal of Finance*, *62*(1), 251–301.

Hochberg, Y. V., Ljungqvist, A., & Lu, Y. (2010). Networking as a barrier to entry and the competitive supply of venture capital. *The Journal of Finance*, *65*(3), 829–859.

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Kashcha, A. (n.d.). *Graph of related subreddits*. Retrieved June 27, 2022, from https://anvaka.github.io/sayit/?query=

Katila, R., Rosenberger, J. D., & Eisenhardt, K. M. (2008). Swimming with Sharks: Technology Ventures, Defense Mechanisms and Corporate Relationships. *Administrative Science Quarterly*, *53*(2), 295–332. https://doi.org/10.2189/asqu.53.2.295

Kerr, W. R., Lerner, J., & Schoar, A. (2014). The Consequences of Entrepreneurial Finance: Evidence from Angel Financings. *The Review of Financial Studies*, *27*(1), 20–55. https://doi.org/10.1093/rfs/hhr098

Kim, W. C., & Mauborgne, R. (2005). Value innovation: A leap into the blue ocean. *Journal of Business Strategy*.

Krackhardt, D. (1988). Predicting with networks: Nonparametric multiple regression analysis of dyadic data. *Social Networks*, *10*(4), 359–381.

Kvalnes, Ø. (2019). *Moral Reasoning at Work: Rethinking Ethics in*

Organizations. Springer International Publishing. https://doi.org/10.1007/978-3-030-15191-1

Lee, S. U., Park, G., & Kang, J. (2018). The double-edged effects of the corporate venture capital unit's structural autonomy on corporate investors' explorative and exploitative innovation. *Journal of Business Research*, *88*, 141–149.

Leiblein, M. J., Reuer, J. J., & Dalsace, F. (2002). Do make or buy decisions matter? The influence of organizational governance on technological performance. *Strategic Management Journal*, *23*(9), 817–833.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal*, *14*(S2), 95–112.

Mallaby, S. (2022). *The power law: Venture capital and the making of the new future*. Penguin Press.

Malthus, T. R. (2007). *An essay on the principle of population*. Dover Publications.

McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61. https://doi.org/10.25080/Majora-92bf1922-00a

Moran, P. (2005). Structural vs. relational embeddedness: Social capital and managerial performance. *Strategic Management Journal*, *26*(12), 1129–1151.

N. Berger, A., & F. Udell, G. (1998). The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle. *Journal of Banking & Finance*, *22*(6), 613–673. https://doi.org/10.1016/S0378-4266(98)00038-7

Nash, J. (1950). Equilibrium points in n-person games. *Proceedings National Academy of Sciences*, *36*, 48–49.

Obstfeld, D. (2005). Social networks, the tertius iungens orientation, and involvement in innovation. *Administrative Science Quarterly*, *50*(1), 100–130.

Oliphant, T. E. (2015). *Guide to NumPy*. Continuum Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2018). *Scikit-learn: Machine Learning in Python* (arXiv:1201.0490). arXiv. http://arxiv.org/abs/1201.0490

Perry-Smith, J. E., & Mannucci, P. V. (2017). From creativity to innovation: The social network drivers of the four phases of the idea journey. *Academy of Management Review*, *42*(1), 53–79.

Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: A resource dependence perspective*. Stanford Business Books.

Pfeffer, J., & Salancik, G. R. (2003). *The external control of organizations: A resource dependence perspective*. Stanford University Press.

Phelps, C. C. (2010). A longitudinal study of the influence of alliance network structure and composition on firm exploratory innovation. *Academy of Management Journal*, *53*(4), 890–913.

Porter, M. (1979). How Competitive Forces Shape Strategy. *Harvard Business Review*.

https://hbr.org/1979/03/how-competitive-forces-shape-strategy

Porter, M. (1996). *What Is Strategy?* https://hbr.org/1996/11/what-is-strategy

Powell, W. (1990). Neither market nor hierarchy: Network forms of organization. *Research in Organizational Behavior*, *12*, 295–336.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 116–145.

Preston, S. L. (2004). *Angel investment groups, networks, and funds*.

Reback, J., McKinney, W., Jbrockmendel, Bossche, J. V. D., Augspurger, T., Cloud, P., Gfyoung, Sinhrks, Klein, A., Roeschke, M., Hawkins, S., Tratner, J., She, C., Ayd, W., Terji Petersen, Garcia, M., Schendel, J., Hayden, A., MomIsBestFriend, … Mortada Mehyar. (2020). *pandas-dev/pandas: Pandas 1.0.3* (v1.0.3) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.3715232

Reuber, A. R., Knight, G. A., Liesch, P. W., & Zhou, L. (2018). International entrepreneurship: The pursuit of entrepreneurial opportunities across national borders. In *Journal of International Business Studies* (Vol. 49, Issue 4, pp. 395–406). Springer.

Rothaermel, F. T., & Deeds, D. L. (2004). Exploration and exploitation alliances in biotechnology: A system of new product development. *Strategic Management Journal*, *25*(3), 201–221.

Schildt, H., Keil, T., & Maula, M. (2012). The temporal effects of relative and firm-level absorptive capacity on interorganizational learning. *Strategic Management Journal*, *33*(10), 1154–1173.

Schumpeter, J. A. (1939). *Business cycles* (Vol. 1). McGraw-Hill New York.

Scott, W. R. (2001). *Institutions and organizations: Ideas, interests, and identities* (Fourth edition). SAGE.

Seabold, S., & Perktold, J. (2010). *Statsmodels: Econometric and Statistical Modeling with Python*. 92–96. https://doi.org/10.25080/Majora-92bf1922-011

Smith, S. W., & Hannigan, T. J. (2015). Swinging for the fences: How do top accelerators impact the trajectories of new ventures. *Druid*, *15*, 15–17.

Smith, S. W., & Shah, S. K. (2013). Do Innovative Users Generate More Useful Insights? An Analysis of Corporate Venture Capital Investments in the Medical Device industry: Do Innovative Users Generate More Useful Insights? *Strategic Entrepreneurship Journal*, *7*(2), 151–167. https://doi.org/10.1002/sej.1152

Stinchcombe, A. L. (1965). Structure and Organizations. *March, J.P., Ed., Handbook of Organtzatlons, Rand McNally, Chicago, 142-193.*

Stock, J. H., & Watson, M. W. (2020). *Introduction to econometrics* (Fourth edition, global edition). Pearson.

Straits, B. C., & Singleton, R. (2018). *Social research: Approaches and fundamentals* (International Sixth Edition). Oxford University Press.

Teece, D. J. (2018). Business models and dynamic capabilities. *Long Range Planning*, *51*(1), 40–49.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*(7), 509–533.

Tiger, L., & Fox, R. (2017). *The Imperial Animal*. Routledge.

https://doi.org/10.4324/9781315132648

Tiwana, A. (2008). Do bridging ties complement strong ties? An empirical examination of alliance ambidexterity. *Strategic Management Journal*, *29*(3), 251–272.

Uzzi, B. (1997). Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness. *Administrative Science Quarterly*, *42*(1), 35. https://doi.org/10.2307/2393808

Uzzi, B. (2020). *Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness*. 34.

Vishwanath, S. R. (2007). *Corporate Finance: Theory and Practice.* SAGE India. http://www.myilibrary.com?id=201766

Von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior* (60th anniversary ed). Princeton University Press.

Wadhwa, A., & Kotha, S. (2006). Knowledge creation through external venturing: Evidence from the telecommunications equipment manufacturing industry. *Academy of Management Journal*, *49*(4), 819–835.

Williamson, O. E. (1991). Comparative Economic Organization: The Analysis of Discrete Structural Alternatives. *Administrative Science Quarterly*, *36*(2), 269. https://doi.org/10.2307/2393356

Wiltbank, R., & Boeker, W. (2007). Returns to angel investors in groups. *Available at SSRN 1028592*.

Winston Smith, S. (2021). Spanning Two Worlds? Corporate Accelerators and Corporate Venture Capital in Innovation Portfolios. *Academy of*

Management Proceedings, *2021*(1), 15786.

Winston Smith, S. W. (2019). Hard to get? Tracking Knowledge Heritage From Invention to Product Market Innovation Using NLP. *Academy of Management Proceedings*, *2019*(1), 18011. https://doi.org/10.5465/AMBPP.2019.18011abstract

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data MIT press. *Cambridge, MA*, *108*(2), 245–254.

Zhang, L., Gupta, A. K., & Hallen, B. L. (2017). The conditional importance of prior ties: A group-level analysis of venture capital syndication. *Academy of Management Journal*, *60*(4), 1360–1386.

Zhang, Y., & Li, H. (2010). Innovation search of new ventures in a technology cluster: The role of ties with service intermediaries. *Strategic Management Journal*, *31*(1), 88–109.