A RISK PERSPECTIVE ON THE RELATION BETWEEN INVESTORS AND THE DIGITAL INFRASTRUCTURE OF STARTUPS

MATTHIAS SCHULTE-ALTHOFF
Freie Universität Berlin
Garystr. 21
14195 Berlin, Germany

KAI SCHEWINA
Freie Universität Berlin

DANIEL FÜRSTENAU
Freie Universität Berlin

INTRODUCTION

The literature on digital entrepreneurship has started to notice the novel nature of the relationship between investors and startups given the needs and opportunities of digital infrastructure (Nambisan, 2016; Jarvenpaa and Markus, 2018). Beyond mere capital streams, some investors provide office space, mentoring (Iansiti and Levien, 2004), strategic thinking (Zahra and Nambisan, 2012), guidance, networking, and other types of services. For this relationship to be successful, we know that different factors play a role such as spatial proximity (Lutz et al., 2013), mutual trust (Bottazzi et al., 2011), social capital of the investor (Aldrich, 2014), as well as passion and other personality characteristics of the involved parties (Mitteness et al., 2012). In addition, we know that these digital startups and their investors form dense entrepreneurial ecosystems to engage in productive entrepreneurship within a geographically defined territory (Basole et al., 2018).

The purpose of this paper is *to better understand the relationship between investors and technology stacks of digital startups in entrepreneurial ecosystems with respect to risk of technological decisions*. By investors, we refer to the type of actor supporting a startup in its early growth phase (i.e. *early stage investors*, such as accelerators or incubators, and *venture capitalists*). We refer to technology stacks (Tech Stacks) as the set of software-based technologies used in these startups and their modular arrangement into different categories. Given the important role that startup firms often play in identifying and commercializing new technologies, we focus on the effect of investment relationships on *technological homogeneity*. While other streams of literature mainly concentrate on ecosystems as sources of competitive edge for companies (Iansiti and Levien, 2004), we will adopt a risk perspective on digital infrastructure as suggested by the digital entrepreneurship literature drawing on an ecosystem point of view (Nambisan, 2016).

Our study makes use of data sets collected from the large-scale web content aggregators Stackshare and Crunchbase. Given the difficulties to observe the relationship between investors and Tech Stacks of digital startups directly, we look into the technological similarity in a network of startups, in which an investor is investing in. We then look into the distinctness of the *technological homogeneity* for different investor types. We find that early stage and venture capitalist startup investors have more homogeneous technology stacks than other types of investors. We argue, these results favor cluster risks of an investor's more homogenous portfolio while at the same time synergy effects may occur.

These findings point to additional complementarities and risks of investors and digital startups beyond the financial realm and previously largely unnoticed consequences of investor choice for startups in the digital age. Thus, we extend the current research on collaboration and cooperation in the entrepreneurial ecosystem with focus on the use of digital infrastructure. Additionally, our "building block" understanding of digital infrastructure as organized in Tech Stacks is potentially valuable with companies becoming more distributed, networked and depending on co-creation while, at the same time, technologies become more layered and modular.

THEORY AND HYPOTHESES

In this section, we present the *technological homogeneity* between startups' Tech Stacks and investors as the outcome variable of interest in our study. We explain the suspected relationships between Tech Stacks of startups and early stage as well as venture capital investors. To do so, we first present the concept of entrepreneurial ecosystems focusing on information about the corresponding sectors, spatial information and investor relationships. We then shine light on the importance of digital infrastructure of startups from an ecosystem point of view closing with two hypotheses bringing together both technical and investment aspects.

Entrepreneurial Ecosystems

Venture capitalist firms are one of the most typical forms of investment, if a conventional business credit via a bank is not available (Gompers and Lerner, 2001). In addition to providing pecuniary support, venture capitalists often give advice, such as identifying business opportunities, potential business partners and networks, developing skills as well as recognizing opportunities on the market (Sapienza et al., 1996). Both types of investors assist companies in different ways often aiding during the creation of the business plan, receiving capital or expert knowledge (Grimaldi and Grandi, 2005). The associations between digital entrepreneurs and investors are presented in this paper in a network-like structure, in which the hubs (investors) try to support the scattered endeavors of the entrepreneurs – sometimes across different industry sectors.

Understanding of Digital Infrastructure as Organized in Tech Stacks

In the following, we want to focus on the intangible technological part of digital infrastructure as described in Lyytinen et al. (2017). We operationalize this concept by focusing on Tech Stacks – software, communication technology, data and core application as being listed by Stackshare. Focusing on the Tech Stack, we can think of the elements as building blocks for the construction of larger components – the foundation of gradually enhancing the business model or the technological basis of your company (Arthur, 2009).

Homogeneity of Tech Stacks as a Measure of Risk in Investors-Startup Relationships

Global networks such as entrepreneurial ecosystems increase in size and complexity. They create risks for individuals as well as companies or states and therefore are in need to be protected (Mertens & Barbian 2014). Because of the unbearable expenditure that would be connected to simultaneous protection of all nodes and edges within the network, the focus should be on the critical elements. Therefore, the challenge is to identify and specifically protect the systemic elements, such as the greater technological interdependencies within the system (Worrel & Bush, 2007).

There is reason to suspect that startups and investors both profit from an exchange of information and investors indeed affect decision-making. The relationship between both parties is of a complex kind functioning in a non-hierarchical way allowing for both tight individual bonds as well as information diffusion (Sapienza and Korsgaard, 1996). It is this diffusion of expertise between investor und investee, which can foster the use of similar Tech Stacks within a cluster in the corresponding entrepreneurial ecosystem. This similarity of the Tech Stack of startups supported by the same investor we want to describe and measure in this paper as the *technological homogeneity* as depicted in figure 1.

Figure 1 about here

In addition to the investment relationship between investor and startups, our analysis considers spatial proximity, funding rounds and the sector-specific diversity of the investor. Of course, there are many more variables, which potentially influence the *technological homogeneity* between investor and investee that we do not consider, such as social capital, startup-hub dynamics and individual relationships – factors that are beyond the scope of this quantitative analysis. Given the aforementioned findings, we posit:

> *Hypothesis 1: The homogeneity of Tech Stacks of startup investors (venture capitalists and early stage investors) is greater compared to other types of investors.*

Adding to this line of thought, the results of Gompers et al. (2009) propose a "strong positive relationship between the degree of specialization by individual venture capitalists at a firm and its success." Thus, it has been argued that more experienced venture capitalists regarding technology and founding experience are superior in selecting better ideas and startups, lowering costs when developing new products and supporting managerial decisions (Breznitz et al., 2018). Therefore, we assume that an investor's specialization in a certain sector approximates for professional experience. We therefore put forward our second hypothesis:

> *Hypothesis 2: The homogeneity of Tech Stacks of startups invested in by startup investors (venture capitalists and early stage investors) is greater if the investor's portfolio is less diverse regarding the corresponding sectors of the investees.*

DATA ANALYSIS AND RESULTS

Our data analysis approach had three main steps: data retrieval, analysis and results interpretation. Phase 1 covered the data extraction via the Crunchbase API and via collecting data from the Stackshare website. This step entailed data preparation and curation to enable proper information processing. Phase 2 consisted firstly of determining the Tech Stack similarity between startups and, secondly of both a variance and a visual network analysis. Finally, in Phase 3, we combined the results to understand better the considered entrepreneurial ecosystem and to test the hypotheses of the last subchapter.

Figure 2 about here

Following the studies of Basole et al. (2016) in the field of visual decision support for ecosystem analysis, it has been indicated that network representation is outperforming other frequently utilized methods as matrices or lists for explorative visualization of complex systems. The network of the entrepreneurial ecosystem considered in this paper depicts investors and startups as nodes and connects investors and its invested startups via edges. Figure 2 shows the resulting network graph. Thirty-nine clusters of high density regarding the similarity of the Tech Stack of the analyzed startups were identified and corresponding nodes are colored accordingly. A force-directed network layout was applied, which arranged nodes based on laws of attraction and repulsion. Finally, an edge weight filter was used to avoid clutter. We chose an edge weight equivalent to the respective similarity indices. A sensitivity analysis as shown in the table on the right in figure 2 has shown that a minimum similarity of 0.05 is sensible: The cluster number rises erratically using a minimum similarity above 0.1. By contrast, a minimum similarity value of 0.05 leads to a high modularity while not producing too many clusters that are devoid of explanatory power. Exemplarily, within a triangle of three big venture capitalists, it can be seen that *Andreessen Horowitz* and *SV Angel* are positioned close to each other, i.e. exhibiting similar Tech Stacks. In contrast, another major venture capitalist, *500 Startups,* was found to be dissimilar to both.

Analysis of Variance

We conducted a Welch's ANOVA test to measure the Tech Stack similarity between startups and investors, which represents our concept of *technological homogeneity*. We consider this variance analysis under different filter conditions, such as only considering a certain sector, funding round or spatial proximity. Mean indices of a present investor relationship and no-investor relationship were compared by investor types and sector categories. We included funding round and spatial proximity as control variables.

First, we found a positive effect of being invested in by one of the considered investors on the *technological homogeneity*. The effect size does not significantly differ when comparing venture capitalists and early stage investors. Both early stage and venture capital investor relationship leads to an increase of almost 50% in similarity in comparison to not being supported by the same investor. Fragmentation by industry sectors show similar results concerning the differences in effect sizes by investor types. The highest effects can be observed in the health sector – a result that may correspond to the low investor entropy we measured. Surprisingly, the funding rounds of the startups did not show a measurable effect on the fit in our model.

Since the ANOVA let us assume that there is a more pronounced *technological homogeneity* due to an investor relation of an early stage investor or a venture capitalist in comparison to the baseline of all investor types, we have assumptions to confirm hypothesis 1. In a similar way, we find indications to assume that sector-specificity leads to a higher homogeneity since the measured results are strongest in the sectors of the lowest investor entropy and therefore that hypothesis 2 is correct.

DISCUSSION AND CONCLUSION

We aimed to better understand the relationship between investors and technology stacks of digital startups in entrepreneurial ecosystems from a risk perspective. Using a broad data set of publicly available data incorporated by information aggregators on the web, we have shown: First, there is a more pronounced homogeneity between investors and the Tech Stack of the invested companies and, second, that this *technological homogeneity* is greater when investors are sector-specifically more specialized. Factors like spatial proximity and the startups funding round influence the measured homogeneity as well but fail to explain the differences in variance in its entirety by themselves.

We contribute to the part of digital entrepreneurship literature dealing with digital ecosystems (Jarvenpaa and Markus, 2018; Basole et al, 2018). We are expand the literature by considering a perspective of technological risk and examining the previously undertheorized relationship between investors and their influence on the technology stack of digital startups. While taking into account sector-specifity and geographical proximity, we find that the Tech Stacks of startups being funded by venture capitalists or early stage investors are more homogenous – an effect even stronger in technologically more specialized sectors as finance or health care.

Our findings provide insights about the denser technological connection regarding startups being funded by the same early stage investor or venture capitalist and give reasons to believe that investors are facing greater cluster risks because of the homogenous technology portfolio. At the same time the question is provoked whether there are tighter bonds within these more akin clusters and if these bonds yield more successful business ventures. After all, Radojevich-Kelley and Hoffman (2012) found that accelerator graduates have higher success rates compared to non-accelerator graduates as measured by longevity in business and receipt of further funding.

We also make a methodical contribution to the understanding of digital infrastructure in digital entrepreneurship contexts (Nambisan, 2016). We provide a data-driven method for operationalizing the technology stack of digital startups as well as investment information based on a unique combination of large amounts of publicly available data. Often, empirical analyses in this field rely on relatively small proprietary datasets or individual case studies limiting its generalization. Through our study, the technology stacks of different digital startups can be compared and similarities and differences can be identified on a broad basis.

REFERENCES AVAILABLE FROM THE AUTHORS

FIGURE 1

Research model on the relationship between an investor and the *technological homogeneity* of Tech Stack of the startups invested in by the investor. The homogeneity measure describes the similarity of digital infrastructure of startups having the same investor.

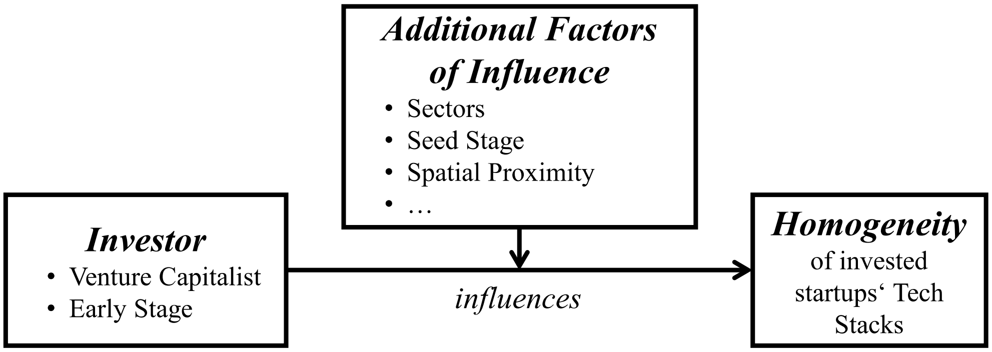
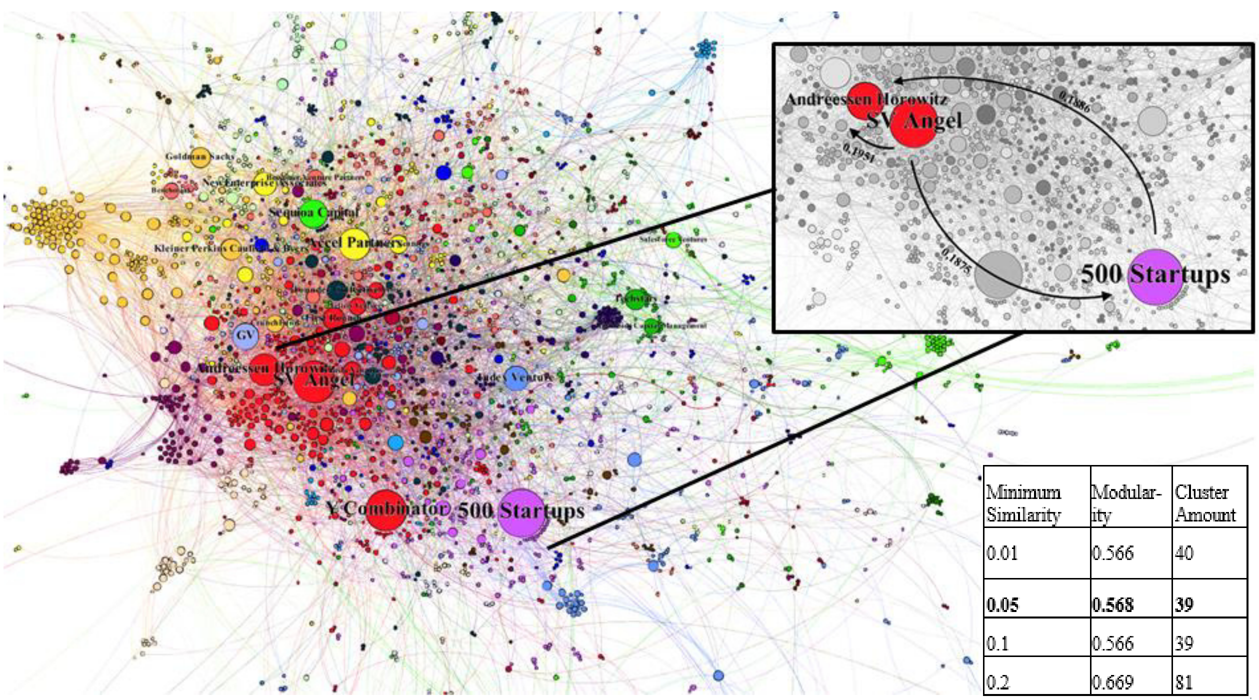


FIGURE 2

Entrepreneurial ecosystem visualized in a network structure. It depicts startups and investors as its nodes and investment relationships between them as its edges. The colored clusters denote the similarity of Tech Stacks of startups respectively of the startups invested in by an investor. At the bottom, a sensitivity analysis regarding the edge weight filter is included, which depicts a rapid increase of clusters starting at a minimum similarity of 0,2. We chose 0,05 as a reasonable value.



Minimum Similarity	Modularity	Cluster Amount
0.01	0.566	40
0.05	**0.568**	**39**
0.1	0.566	39
0.2	0.669	81