Finding the Next Unicorn: When Big Data Meets Venture Capital

Johannes Weibl
LMU Munich
weibl@bwl.lmu.de

Thomas Hess
LMU Munich
thess@bwl.lmu.de

Abstract

Venture capital (VC) has been growing rapidly in recent years. So far the screening and evaluation of potential startups as investment objects largely depends on the venture capitalist's personal experience, network and qualitative evaluations. In the era of big data, the advent of new data sources and analytic techniques enables a data-driven investment process. Grounded in systems theory and the theory of complementarity, this study reports the findings from an exploratory study of 13 VC firms that synthesize and use novel data sources. Our analysis shows that the data-driven approach, in particular, impacts the deal origination and screening stages of investment. It leads to informational and transactional benefits, which lower operational costs in the short term and enlarge the potential return on investment of a VC firm in the long term. We contribute to the literature by shedding light on how various data sources complementarily lead to additional business value.

1. Introduction

In the few last years, the economy in Europe and the US has experienced a significant boom in venture capital (VC). This phenomenon is mainly driven by the "explosion of startups" and the maturity of the entrepreneurial finance market [1]. In the US, for instance, venture capitalists (VCist) invested $74.2 bn. in 5268 deals in 2017 according to PwC [2]. In Germany in 2016, the top 100 startups received funding in the amount of $5.9 bn. [3].

By offering mentoring and financial resources, VC firms can receive high return on their investments if their portfolio companies are being acquired by another company or going the route of an initial public offering (IPO). VC refers to the financing of private companies and is a form of private equity, medium to long-term finance provided in return for an equity stake in potentially high growth companies [4]. By being the biggest capital supplier, VC firms support entrepreneurial firms to facilitate, for example, their operations, production, or expansion [1].

Up until now, the screening and assessment of startups for investments often relied on the investor's experience, relationships and networks, as well as on qualitative assessments of entrepreneurial firms [5]. However, in the era of big data, more accessible data sources and data explosion as well as more sophisticated analytic techniques can facilitate a more data-driven decision-making and investment process for VC firms.

Venture capital has been discussed variously in the academic literature. Prior studies have investigated the performance of VC investments regarding value creation and innovation power [6, 7], various investment strategies [8], as well as recommendation systems in the domain of venture finance [4]. Other studies shed light on data analytics and data-driven models in the VC context, for instance, to enhance the existing prediction model to ensure better investment decisions [9].

However, it still remains unclear how the broad availability of new data sources and analytic technologies (often termed as big data [10, 11]) are used by VCists during the investment process and how additional business value is generated. Thus, this paper addresses the following research question: *How does big data transform the investment process of VC firms?*

In order to answer this question, we build on the literature of big data analytics, systems theory [12] and the theory of complementarity as a theoretical foundation. Methodologically, we conducted interviews with 13 VCists in Germany with the objective to reveal the impact of using various data sources within their investment processes.

This research makes important contributions to both theory and practice. First, we contribute to prior research by providing a qualitative empirical study in the field of systems theory and the theory of complementarity. Second, our findings indicate that VC firms synthesize their existing transactional IT systems, e.g., a customer relationship management (CRM) system, with external, web-based data

HICSS

services for better insights and data-driven decision making. The advent and usage of more information transforms, in particular, deal origination and screening, stages 1 and 2 respectively of the five-stage investment process. It leads to an automation of sourcing and screening of entrepreneurial firms and substantially enhances the ability to identify successful investment deals. In the short term, the data-driven transformation of VC firms leads to lower operational costs, and in the long term the usage of data analytics leads to greater return on investments.

Third, this study reveals the path of value generation through data and data analytics in the era of big data. For VC firms it provides clear guidance how to utilize the additional value of their investments in a data-driven environment by synthesizing their existing data systems with external web-based services.

This paper is organized as follows: In section 2, we present the study's theoretical background (systems theory and investment process). We then describe the research methodology (section 3) and present and discuss our findings (section 4). The paper concludes by highlighting the research and practical contributions that our paper makes (section 5).

2. Theoretical background

2.1. Big data sources and analytics

Big data as a term is generating significant attention in a range of companies worldwide. It refers to the process of managing large amounts of data that come from heterogeneous data sources (e.g., internal and external, structured and unstructured) that can be used for collecting and analyzing an enterprise's data for the purpose of augmenting its performance [13]. The phenomenon "big data" is often summarized by the notion of three Vs: volume, velocity and variety. It refers to the large volume of data that companies store and analyze nowadays, generated from a greater variety of data sources with multidimensional data fields with a high frequency of data generation and data delivery [14]. In this paper we follow Davenport et al. [10] and Duan and Cao [11] who focus on the variety of data sources within the big data context. Furthermore Goes [15] and Malgonde and Bhattacherjee [16] highlight the importance of data from various data sources as the major driver for additional business value in the era of big data. In other words, big data and related analytics (BDA) is about the extraction of unknown patterns,

correlations and information across multiple sources of data to enrich the information depths and produce new insights for decision makers [11].

Following Chen et al. [13], we identify three groups of data sources for companies (see Table 1). Driven by the diffusion of new technologies (such as mobile devices and social media platforms), advancements and new developments of storage and integration possibilities (e.g., data lake), organizations are able to synthesize internal company data sources (e.g., across divisions) as well as integrate external data sources.

Table 1. Groups of data sources [13]

Groups	Description
Transactional systems	Data from enterprise systems, including standard workflow and automation tools, such as CRM systems.
Web-based services	Data from online systems, such as e-commerce platforms and social media systems.
Mobile systems	Data from sensors (e.g., location data) generated from mobile devices.

2.2. Systems theory and business value

From a theoretical perspective, our study is based on systems theory, which looks at systems as a whole rather than at its individual parts [17]. It addresses the interplay of resources in organizations that collectively create greater value than any of the resources in isolation [18]. A system's individual components are described as subsystems [19]. If different subsystems (a system's individual components [19]) interact complementarily, the overall value of a system can exceed the value of each subsystem in isolation. Complementarity, defined as the ability of one resource to reinforce the impact of another resource, is often described as the underlying activity leading to enhancing effects on the relationships between resources and outcomes and is derived from the economic theory of complementarities [20].

Data sources can be understood as IT resources in companies with the ability to share or aggregate information [21]. According to systems theory, each data source can be conceptualized as a data subsystem in a firm. Data subsystems can be complementary to each other [22] allowing actors deeper insights from their combination [23]. Combined data and related data sources have a greater value than the sum of its individual parts, which facilitates decision makers to derive more

insights and leading to a better decision-making process [24]. When various data sources are combined, data complementarily interact with each other resulting in additional business value [22].

In general, the complementary interaction of different IT systems generates a wide range of both tangible and intangible business value [25]. It can be realized and measured in three dimensions: transactional, informational and strategic [26, 27]. Strategic and transactional benefits can be seen as tangible benefits because they can often be measured in e.g., productivity, savings, innovation or market share [28]. Transactional benefits include effectiveness, process efficiencies as well as automation and are often associated with cost reductions [26]. Strategic benefits include increased revenue, superior customer experience, shorter time to market, and development of new products and services [27]. Informational benefits as intangible benefits, such as enabling, improving, increasing, etc., are more difficult to measure. Informational benefits also include improved information for managing, enhanced reporting and communication with customers and suppliers. Furthermore, they support the responsiveness and reliability of firms and enable more effective decision making [26]. Moreover, informational benefits can lead to real-time decisions and new pattern discovery [27].

2.3. Venture capital and the investment process

Venture capital, defined as a form of private equity invested in high growth companies in return for an equity stake is often seen as an extreme form of financing [4]. Previous studied in the finance and corporate governance literature have analyzed the role for VCs in resolving various agency problems, which arise in multistage financial contracting problems [29], identified economic and legal determinants of VC investments and exits [30], and studied the effects of various investment strategies (e.g., pure-play versus more general investment strategy to minimize portfolio risks) [31]. Yet, regarding the usage of new data sources and capturing related business value in VCs, little research exist, although the impact of big data on companies can be large [13].

With our research question in mind, we shed light on an investment model that explains the investment activities of VCists. Several researchers have discussed various process models, amongst the most cited is that of Tybjee and Bruno [5], who developed a five-stage investment model (see Table 2). Tybjee

and Bruno's model helps to analyze and structure our findings regarding the impact of big data on the investment process.

Table 2. Venture capital investment process [5]

Stage	Description of VC firm activity
Deal origination	Becoming aware of potential investment activities.
Screening	Concentrating only on a manageable set of potential deals.
Evaluation	Assessing the potential return and risk of a particular deal. If the outcome of the evaluation is positive, the VC firm enters into negotiations with the potential investee.
Structuring	Structuring the deal in terms of the amount, form and price of the investment.
Post-investment activities	Setting up controls to protect the investment, providing consultation to the fledgling management of the venture and supporting the orchestration of the acquisition.

3. Research method

The goal of this study is to explore how the usage of data and novel data sources impacts the investment process of VC firms. To elaborate the impact in detail, we decided to use qualitative research with an explorative design. The qualitative approach enabled us to analyze in detail the impact and relationship between different factors and to consider contextual factors [32].

In IS research, exploratory qualitative study has been a legitimate means of research. Because of the complexity of our chosen subject area and our goal to explore the investment process and potential impact in detail, this research approach is appropriate. Furthermore, it stimulates further research questions for related work. In fact, the explorative approach allowed us to analyze our research subject in areas in which only limited knowledge exists [32, 33]. To achieve our goals, we conducted semi-structured interviews.

3.1. Data collection

Data collection took place in 13 different organizations in the VC sector in Germany. Analogous to Islam et al. [34], we focused on the primary key informant at each firm as our research subject. In total, we conducted 13 semi-structured

interviews (telephone: 10; face-to-face: 3) with technology and big data experts at VCs. The interviewees were mostly on the partner level of the firm or were employees in key decision making roles (e.g., director, principal, investment manager). We chose employees of both corporate and independent VC firms with various investment foci to present an industry-specific picture of the data-driven impact on the firms. We brought new interview experts continually into our study until we received a high rate of duplications of responses. After our 13 interviews we reached this point of saturation. So, we did not extend the interview process further but started to analyze the interview data. Table 3 provides an overview of the data collection.

Table 3. Overview of data collection

#	Position of interviewee	Type of VC	Investment Focus
1	Associate	Independent	Digital health firms
2	Director	Independent	Travel and consumer goods firms
3	Analyst	Corporate	Energy startups
4	Director	Corporate	Mobility and energy startups
5	Partner	Independent	Technology firms
6	Associate	Independent	Software-as-a-service firms
7	Investment Manager	Independent	Digital startups
8	General Partner	Corporate	Software-as-a-service startups
9	Analyst	Independent	Broad focus
10	Data Scientist	Independent	Digital startups
11	Principal	Independent	Technology firms
12	Associate	Corporate	Energy startups
13	Management Director	Independent	Technology focus

Prior to conducting the interviews, we prepared an interview guideline following Schultze and Avital's guide [35]. The interview guideline was based on literature and included three parts: 1) general aspects of the VC firms (investment focus, employees, etc.), 2) the firms' usage of internal and external data-driven IT systems and 3) their usage of data-driven IT systems for each investment stage and its potential business value. Depending on the respective knowledge of the interviewees, we asked additional questions. The interviews lasted between 45 and 80 minutes and were recorded, transcribed and anonymized using the software ATLAS.ti. In addition to our interviews, publicly available data were used to triangulate our findings (press releases, management reports, VC websites, etc.)

3.2. Data analysis

The goal of our data analysis was to glean essential insights by abstracting a manageable collection of interview data, while still reflecting the interviews. This methodological technique is known as content analysis [36]. Our coding was based on the descriptive coding scheme derived from the stages and impact of the investment process model and underlying technologies. The coding was conducted independently by two researchers to ensure quality data analysis. Both researchers read all the transcripts in preparation for the coding, then familiarized themselves with the interview data guided by the fundamental principles of the hermeneutic circle [37]. After completing the initial coding, we used a consensual approach to reduce obvious differences between the researchers and to refine the coding scheme.

4. Insights

In this section, we describe our results and discuss the insights from our qualitative research approach. First we present our insights on data and relevant data systems of VC firms, second, we show their impacts on the investment process and third, we describe its business value.

4.1. Venture capital-relevant data and related systems

Our interviews reveal that VC firms use both transactional IT systems and web-based data services as groups of data sources to leverage their investment activities. The transactional system, often in the form of a CRM system, is the backbone of the VC IT system. "*[It is used] to track our deal flows and contacts in the firm,*" one interviewee stated. In the VC context, it is also used to log meetings, maintain client contact details and structure marketing and sales leads. The web-based data services include various types of external data and data-driven service providers.

Data from both types of systems interact in two ways: a) data from web-based services enrich

existing, internal data from transactional systems and b) various data within web-based services are combined within the external platform for further insights for the VCist. We interpret this finding as two types of complementary interaction (see Figure 1).
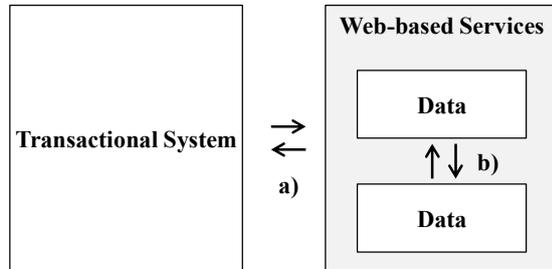

Figure 1. Two types of complementary interaction

Based on our interviews, we found that VC firms typically integrate social media data, such as contact information and related activities, into their internal VC CRM systems. This makes it possible to "*reach people more easily and faster*" according to many interviewees and provides further insights about the general opinions (e.g., political opinions) and preferences of entrepreneurs, which, in turn, may influence the investment decision. Many of the VC firms also synthesize external market and customer information from market intelligence platforms (e.g., Mattermark, Statista, Gartner) with their internal market information. This sheds light on potential consumer trends and new market opportunities, which cannot be validated without external market information. "*By merging information from Statista and Mattermark into our system we are able to assess whether the market conditions for a specific startup are rather favorable or not*" (The director of VC firm #2).

A further insight emerged from our interviews when we investigated the usage of web-based services: VCists typically have access to various crowdfunding websites and investment networks. Here the objective is not necessarily to integrate these data sources into their internal systems, rather, VCists analyze and combine data directly on the crowdfunding sites or investment networks for further insights. Analyzing data from crowdfunding sources helps to keep track of new deals in the VC sector and innovative products and services that are being started or generating significant traction. Additionally, investment network services allow for insights on team structure, team expertise, money raised up to now, the success or failure of previous startups of the entrepreneurs, etc. Combining these

insights helps to further identify high potential entrepreneurs and startups and reduces the risk of investment failures from a VC perspective. One interviewee mentioned that by synthesizing data from Crunchbase and Pitchbook, they found out that an investment deal with founders from various universities was more likely to succeed than one whose founders came from the same university. Moreover, web-based benchmarking services allow VCists to evaluate their portfolio companies against relevant peer groups in e.g., benchmarking financial performance.

As an interim result, Table 4 summarizes the data and VC related systems.

Table 4. Overview of venture capital-relevant web-based services

	Groups of Data Sources	Data (excerpt)
a) Synthesize data with trans-actional systems	Social media platforms, e.g., LinkedIn	Contact information, social graph, founders' opinions
	Market intelligence platform, e.g., Mattermark, Statista	Consumer trends, market opportunities
b) Combining data within web-based service	Crowdfunding sites & blogs, e.g., Producthunt	New deals, product and company launching
	Investment networks, e.g., Crunchbase, Pitchbook	Team and related expertise, money raised
	Benchmarking platforms, e.g., Social Capital	Business data from market and competitors

4.2. Impact on the investment process

Based on the insights from the previous section that VC firms use primarily transactional data systems and web-based data services, we show in this section how data usage impacts each of the five investment stages [5].

Deal origination: In the beginning stages of any investment, VCists face the challenge of identifying prospective investment deals. Our interviews reveal that in the past, next to typical digital inbound services (such as emails and company website forms), various intermediaries (such as lawyers)

played an important role in shedding light on potential startup companies: "*Mostly it was a combination of different sources of information, personal networks, personal recommendations, which then come, for example, also from other VCists or attorneys*" (Director of VC firm #4).

However, the deal origination process fundamentally transforms from an inbound orientation to an outbound orientation by means of the usage of data. By accessing investment networks such as Pitchbook or Crunchbase, VCists combine company information with market research for a better overview of newly founded startups, which are potential investment deals. Typical crowdfunding websites (e.g., Producthunt) further support the deal origination of VCists by providing data about the existence of new crowdfunded startups and underlying products or service ideas. Nowadays, VCists have access to market intelligence platforms, e.g., Mattermark, for data-driven insights on market dynamics and related participants. By combining information from various sources, such as website traffic, employee counts, time since last funding, co-investors and total amount of funding, these services allow VCists to predict the growth and success of a startup. Some interviewees further stated that their firms are "*automatically informed about the existence of a new company. We use customized data tools, sometimes including self-learning algorithms to find investments by synthesizing external data sources*" (Associate of VC firm #1).

We conclude that the usage of data, especially from web-based services, has a fundamental impact on deal origination. It reduces the effort required to search for new investment deals and simultaneously provides a broader and more sophisticated basis for decision making.

Screening: During the screening stage, VC firms screen a large number of potential deals based on pre-defined criteria, concerning, for instance, the technology, product or market of the startup. In the pre-big data era this step was less data-driven. As one interviewee stated: "*The screening stage was less data-driven, rather, we focused on qualitative assessments. Discussions and talks with the founders were more crucial for us*" (Investment Manager, VC firm #10).

Nevertheless, our interviews reveal that the screening stage rapidly moves to a more data-driven activity. Here the integration of market intelligence data and social media data in the internal transactional systems (CRM system) plays an important role. Market growth, based on Mattermark or Capital IQ for instance, allows for analyzing and refining business cases as well as return- and risk

models of potential startups in more detail, including a better validation of market growth. As one interviewee stated: "*Everything that goes in the direction of evaluation and business case building is of course strongly data driven, actually all the parts we look at there, the market, the competitive situation, modeling, return and risk modeling, so this is all as quantified as possible. And some of the data is available internally, others we obtain from external providers*" (Director of VC firm #4).

Furthermore, the seamless integration of contact and team information from social media, e.g., LinkedIn, helps VCists in the screening stage to keep in contact with entrepreneurs more easily on various digital channels. Further information provided by social media platforms, such as former startup experiences of entrepreneurs and their academic background, combined with data from investment networks (e.g., team structure, money raised) allows a data-driven approach to determine the probability of a startup's success.

In sum, data usage plays an important role in the screening stage. All the VCist we interviewed mentioned in particular the importance of combining data from external web-based services with the internal CRM system.

Evaluation: Due to the paucity of, or rather non-existence of, operating histories of startups, VC firms have to rely on a subjective assessment based on the business plan and related information delivered by the venture's management to determine the future viability of the startup. Typically, this step in the evaluation aims to subjectively assess the venture using a multidimensional set of criteria, such as management commitment and marketing skills [5].

We recognize that this step is almost always data driven. However, the focus has been mainly on qualitative data, e.g., expert interviews or personality tests, to better rate a potential investment activity. As one interviewee stated: "*In this process we are extremely data driven, i.e., we talk to many industry and technology experts in the respective area. We take a very close look at the market, we examine the business case very closely, and in some cases, we also conduct certain personality tests with the founders*" (Director of VC firm #4).

Furthermore, this step varies substantially from investment case to investment case and heavily relies on the investment team's subjective assessment of the startup. Besides some internal quantitative data that is used for the evaluation, our interviews show that all VC firms use only qualitative information for the final evaluation of a deal.

Structuring: During this step, the investment deal is structured in detail to reach a mutual

acceptable investment agreement between the entrepreneur and the VC firm that includes the price of the deal and further contract details (e.g., management salaries). Our interviews reveal that this step is less data driven overall than manually determined on a case by case basis. "*We do not use any templates or any data-driven standardized process. This may sound odd, but I think that this is actually the case with many VCists, because every deal is different*" (Associate of VC firm #1).

Post-investment activities: Once an investment deal between the VC firm and a startup has been realized, the role of the VC firm transforms from being a pure investor to a being collaborator. Hereafter, the VC firm plays a new role as a formal representative on the board of directors of the startup, and it exchanges market and supplier information informally with the startup.

In the pre-big data era, VCists typically gathered internal data from various portfolio companies to compare a startup performance vis-à-vis its peer-group. "*However, what we are very good at compared to the startup is benchmarks of how similar companies have performed in a similar business phase. For example, if you look at an app, it would be relatively easy for us to ask what the acquisition costs were like or the conversion rate*" (Associate of VC firm #1).

Today, VCists access and use web-based benchmarking tools to compare the competition and peer group of the startup with a broader data base. This helps to aggregate performance data for similar companies and leads to more balanced business decisions. Some VC firms are more progressive and create their own online benchmarking platform during the post-investment stage. The idea behind this is that both portfolio and external startups can upload their core business data for an in-depth benchmark with competitors and derive valuable insights for further growth. Performance data, such as the number of customers, churn rate, conversion rate and fixed costs, are then collected from portfolio companies as well as external startups and synthesized, leading to an enormous amount of valuable data. "*We are working on developing an auto benchmarking service similar to social capital in the USA. They offer such an auto benchmarking service, they say upload your data and then we benchmark that for you, the cool thing is of course that they build up a huge data asset because everything that gets uploaded, save it of course, so I would also be super happy long term*" (Partner at VC firm #5).

Nevertheless, the data-driven benchmarking activities are a sideline activity during the post-investment stage; this step is still characterized by manual and individual management tasks.

Table 5 sums up our findings about the usage of data sources for each investment stage.

Table 5. Usage of data for each investment stage

Investment Stage	Data Usage	Groups of Data Sources
Deal origination	High	Web-based services - Crowdfunding data - Investment network data - Market intelligence data
Screening	High	Web-based services - Social media data - Market intelligence data - Investment network data Transactional systems - Customer and deal flow data
Evaluation	Low	None
Structuring	Low	None
Post-inv. activities	Low	Web-based services - Benchmarking data

4.3. Generating value with data

The usage and combination of transactional and web-based data systems generate a wide range of both tangible and intangible business value for VC firms. We see that the majority of these benefits can be classified as informational or transactional benefits, which especially impact the first two investment stages, deal origination and screening (see Table 6).

By adding further information regarding a potential startup or its market situation from external, web-based services, the VC firm is able to check a larger number of potential startups on a broad data basis in a shorter period of time. The usage of external web-based services, in addition, enables a broader discovery of business models currently emerging and respective products and services in the startup sector, This, in turn, allows for more fact-based decisions about potential investments. Typically, the data-driven approach during the deal origination and screening stages leads to efficiency gains in the short term and greater return on investments in the long term. As one interviewee stated: "*In the beginning it is mainly efficiency gains and in the end it will lead to better returns through better investment decisions*" (Partner at VC firm #5).

The data-driven transformation from an inbound orientation (e.g., receiving e-mails from

entrepreneurs) to an outbound orientation (e.g., actively scanning and analyzing new companies online) leads to an automation of the deal origination and may significantly reduce the time of work needed to identify promising startups. "*Typically, there are not enough resources in VC firms to do everything. Algorithms are a sensible investment for VC firms to improve this ratio of a lot of work and few resources in the long term*" (Associate of VC firm #1).

We see further that a faster, and especially an automated, deal origination and screening process facilitates a shorter time for a VC firm to build the initial relationship with an entrepreneur (summarized as 'time to contact'). As one interviewee stated: "*I think speed is the main advantage, because in the very competitive tech investment sector it is important to talk to companies as early as possible and to be the first to get the chance to establish a relationship with the founders*" (Principal at VC firm #11). We further determine that the informational and

transactional benefits are consecutive and sequential. The usage of various data and data sources leads to informational benefits, followed by a higher degree of automation and efficiency in searching for potential investments and contacting entrepreneurs (transactional benefits).

In addition to the primary outcomes (informational and transactional benefits), our interviews also revealed a potential strategic benefit. Some VC firms not only synthesize external data to gain better insights and improve decision making but also focus on building their own data platforms (e.g., benchmarking platform), which portfolio companies and external startups can use to compare company performance with peer groups. In the long term, this new data source leads to valuable data assets with in-depth information about the VC sector overall.

Table 6. Generating value with data in venture capital

Business Value	Purpose	Affected investment stage	Expected performance benefits
1) Informational	• Discover emerging business models and related products and services to invest in respective products to invest in • Identify entrepreneurs with high potential more quickly and more accurately • Screen potential deals in (nearly) real-time	Deal origination Screening	• Lower costs (short term) • Greater return on investments (long term)
2) Transactional	• Automate deal origination process • Increase the volume of startups identified and screened • Reduce time until initial contact with entrepreneurs (time to contact)	Deal origination Screening	• Lower costs (short term)
3) Strategic	• Provide new services (e.g., benchmarking platform)	Post-investment activities	• Greater return on investments (long term)

5. Contribution, limitations and further research

The key findings from this research involve the observation that many VC firms synthesize their existing transactional IT system, e.g., CRM system, with external, web-based data services to gain better insights and improve data-driven decision making. Various data from web-based data services 1) complement the explanatory information power within the external platform and 2) interact complementarily with internal data from transactional systems. The insights derived from such synthesis and the usage of more information transform, in particular, the first and

second investment stages (deal origination and screening) of the five-stage investment process. They lead to an automation of sourcing and screening of entrepreneurial firms and significantly enhance the likelihood of identifying successful investment deals. In the short term, the data-driven transformation of VC firms leads to lower operational costs, and in the long term the usage of data analytics leads to a greater return on investments.

5.1. Theoretical and practical contribution

From an academic perspective, we contribute to prior research by providing a qualitative, empirical

study based on systems theory and the theory of complementarity. In line with previous literature [e.g., 38], we see that transactional IT systems are implemented mainly to automate processes (e.g., VC deal flow, selection process of new startups) while web-based services as a type of informational system are implemented to inform the same process. Both types together facilitate more intelligent and efficient investment processes, complementing and reinforcing each other. Through this, VCists are able to accomplish higher-order organizational tasks.

Our study also sheds light on the business value emerged from data variety [15]. Several researchers have proposed that value is created from synthesizing various data sources [e.g., 16, 23] in the era of big data. To the best of our knowledge, none of these studies focus on the complementary interaction among various data sources as a theoretical foundation. In our study the complementary interaction helps to understand how the usage and synthesis of various data assets lead to additional business value in organizations.

Furthermore, taking the findings of Wixom et al. [27], it can be argued that informational benefits play an intermediating role in the complementary mechanisms of value creation through data. Initially, the interaction of various data leads to informational benefits for VC firms followed by transactional benefits (e.g., a higher degree of automation and efficiency) and occasionally by strategic benefits (e.g., the creation of new data services). Our study extends existing literature [e.g., 27] by providing an example of sequential logic. Moreover, our study reveals 'time to contact' as a new and additional transactional benefit which has not, to the best of our knowledge, previously received appropriate attention in the current literature [e.g., 26, 27, 39]

In terms of practical contribution, our study shows that advanced analytics and the usage of a wide range of data sources pays off for VC firms. In the short term, enormous gains in efficiency can be realized by the data-driven automation of single investment stages. Practitioners should primarily concentrate on the first two stages (deal origination and screening), where a faster and more detailed evaluation of potential investments can easily be leveraged. The benefits identified provide a solid decision-making basis for triggering expensive data analytic investments within VC firms. This study also shows that a more data-driven way of working should not be a side project in VC firms, but rather should be supported by the entire organization (including top management commitment [e.g., 40]). Establishing the role of a technical partner or chief technology officer (CTO) would effectively signal this.

5.2. Limitations and further research

We acknowledge that our study has some limitations. First, our study was conducted using a qualitative research approach resulting in a limited generalizability of our results. Our insights rely on the experience and opinions of a limited number of interviewees. Although the number of VC firms we engaged is comparable to other VC studies [e.g., 39], we must be careful when generalizing the results. A larger sample size or in-depths case studies of VC firms would provide further insights. Future research could involve more in-depth empirical approaches, such as quantitative analysis of the interview data and further research methods to deepen the insights regarding the derived business value.

Second, we conducted our study with VC firms in Germany. There may be other VC firms, especially in the US and Israel, that are more progressive in applying innovative data-driven projects. Future research could investigate whether these companies have put data and novel data sources to other uses, and thereby shed light on the differences. Further research could also extend the derived business value by in-depth analysis of how VCs and their employees utilize the data and related technologies.

Future empirical studies can take our insights as a fruitful starting point for further research on how the complementary interaction of various data sources impacts business processes, such as the investment process, and validate, as well as revise, our results. To deepen the understanding, our paper should be extended in further research by elaborating in detail the characteristics of the used data sources (e.g., volumes, types of data etc.), which VCs use.

6. References

[1] Zhong, H., Liu, C., Zhong, J., and Xiong, H., "Which startup to invest in: a personalized portfolio strategy", Annals of Operations Research, 263(1-2), 2018, pp. 339-360.
[2] Pwc, "MoneyTree Report Q1/2018": PwC / CB Insights, 2018 pp. 1-79.
[3] Prüver, T., and Greif, M., "Venture Capital and Start-ups in Germany 2016", EYGM Limited, 2017 pp. 1-44.
[4] Stone, T., Zhang, W., and Zhao, X., "An empirical study of top-n recommendation for venture finance", Proceedings of the 22nd ACM international conference on information & knowledge management, 2013, pp. 1865-1868.
[5] Tyebjee, T.T., and Bruno, A.V., "A model of venture capitalist investment activity", Management Science, 30(9), 1984, pp. 1051-1066.
[6] Dushnitsky, G., and Lenox, M.J., "When do firms undertake R&D by investing in new ventures?", Strategic Management Journal, 26(10), 2005, pp. 947-965.

[7] Park, H.D., and Steensma, H.K., "When does corporate venture capital add value for new ventures?", Strategic Management Journal, 33(1), 2012, pp. 1-22.

[8] Yang, Y., Narayanan, V., and Zahra, S., "Developing the selection and valuation capabilities through learning: The case of corporate venture capital", Journal of Business Venturing, 24(3), 2009, pp. 261-273.

[9] Xu, R., Chen, H., and Zhao, L.J., "Predicting Corporate Venture Capital Investment", Proceedings of the 38th International Conference on Information Systems, 2017, pp. 1-10.

[10] Davenport, T.H., Barth, P., and Bean, R., "How big data is different", MIT Sloan Management Review, 54(1), 2012, pp. 43.

[11] Duan, Y., and Cao, G., "Understanding the Impact of Business Analytics on Innovation", Proceedings of the 23rd European Conference on Information Systems, 2015, pp. 1-16.

[12] Ackoff, R.L., "Towards a system of systems concepts", Management Science, 17(11), 1971, pp. 661-671.

[13] Chen, H., Chiang, R.H., and Storey, V.C., "Business Intelligence and Analytics: From Big Data to Big Impact", MIS Quarterly, 36(4), 2012, pp. 1165-1188.

[14] Wamba, S.F., Akter, S., Edwards, A., Chopin, G., and Gnanzou, D., "How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study", International Journal of Production Economics, 165(2015, pp. 234-246.

[15] Goes, P.B., "Editor's comments: big data and IS research", MIS Quarterly, 38(3), 2014, pp. iii-viii.

[16] Malgonde, O., and Bhattacherjee, A., "Innovating using big data: A social capital perspective", Proceedings of the 20th Americas Conference on Information Systems, 2014, pp. 1-9.

[17] Burton-Jones, A., Mclean, E.R., and Monod, E., "Theoretical perspectives in IS research: from variance and process to conceptual latitude and conceptual fit", European Journal of Information Systems, 24(6), 2015, pp. 664-679.

[18] Someh, I.A., and Shanks, G., "The role of synergy in achieving value from business analytics systems", Proceedings of the 34th International Conference on Information Systems, 2013, pp. 1-16.

[19] Nevo, S., and Wade, M.R., "The formation and value of IT-enabled resources: antecedents and consequences of synergistic relationships", MIS Quarterly, 34(1), 2010, pp. 163-183.

[20] Milgrom, P., and Roberts, J., "Complementarities and fit strategy, structure, and organizational change in manufacturing", Journal of Accounting and Economics, 19(2), 1995, pp. 179-208.

[21] Goodhue, D.L., Wybo, M.D., and Kirsch, L.J., "The Impact of Data Integration on the Costs and Benefits of Information Systems", MIS Quarterly, 16(3), 1992, pp. 293-311.

[22] Müller, O., Fay, M., and Vom Brocke, J., "The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics", Journal of Management Information Systems, 35(2), 2018, pp. 488-509.

[23] Günther, W.A., Mehrizi, M.H.R., Huysman, M., and Feldberg, F., "Debating Big Data: A Literature Review on Realizing Value from Big Data", The Journal of Strategic Information Systems, 26(3), 2017, pp. 191-209.

[24] Arisa, S., and D., G.R., "Towards an understanding of the role of business intelligence systems in organisational knowing", Information Systems Journal, 26(4), 2016, pp. 339-367.

[25] Asadi Someh, I., and Shanks, G.G., "How Business Analytics Systems Provide Benefits and Contribute to Firm Performance?", Proceedings of 23rd European Conference on Information Systems, 2015, pp. 1-17.

[26] Aral, S., and Weill, P., "IT assets, organizational capabilities, and firm performance: How resource allocations and organizational differences explain performance variation", Organization Science, 18(5), 2007, pp. 763-780.

[27] Wixom, B.H., Yen, B., and Relich, M., "Maximizing Value from Business Analytics", MIS Quarterly Executive, 12(2), 2013, pp. 111-123.

[28] Mirani, R., and Lederer, A.L., "An instrument for assessing the organizational benefits of IS projects", Decision Sciences, 29(4), 1998, pp. 803-838.

[29] Admati, A.R., and Pfleiderer, P., "Robust financial contracting and the role of venture capitalists", The Journal of Finance, 49(2), 1994, pp. 371-402.

[30] Armour, J., and Cumming, D., "The legislative road to Silicon Valley", Oxford Economic Papers, 58(4), 2006, pp. 596-635.

[31] Knill, A., "Should venture capitalists put all their eggs in one basket? Diversification versus pure-play strategies in Venture Capital", Financial Management, 38(3), 2009, pp. 441-486.

[32] Yin, R.K., Case study research and applications: Design and methods, Sage Publications, 6th, Thousand Oaks, CA, 2017.

[33] Bogner, A., Littig, B., and Mens, W., Interviewing Experts, Palgrave Macmillan UK, Basingstoke, England, 2009.

[34] Islam, N., Trautmann, K., and Buxmann, P., "Tradition Meets Modernity–Learning from Start-ups as a Chance to Create Digital Innovation in Corporations", Proceedings of the 37th International Conference on Information Systems, 2016, pp. 1-11.

[35] Schultze, U., and Avital, M., "Designing interviews to generate rich data for information systems research", Information and Organization, 21(1), 2011, pp. 1-16.

[36] Weber, R.P., Basic content analysis, Sage, 1990.

[37] Klein, H.K., and Myers, M.D., "A set of principles for conducting and evaluating interpretive field studies in information systems", MIS Quarterly, 1999, pp. 67-93.

[38] Zuboff, S., In the age of the smart machine: The future of work and power, Basic Books, New York, 1988.

[39] Fried, V.H., and Hisrich, R.D., "Toward a model of venture capital investment decision making", Financial Management, 1994, pp. 28-37.

[40] Weibl, J., and Hess, T., "Success or Failure of Big Data: Insights of Managerial Challenges from a Technology Assimilation Perspective", Proceedings of the Multikonferenz Wirtschaftsinformatik (MKWI), 2018, pp. 12-59.